
David Mumford

Archive for Reprints, Notes, Talks, and Blog

Professor Emeritus
Brown and Harvard Universities
David_Mumford@brown.edu

[HOME](#) [ALGEBRAIC
GEOMETRY](#) [VISION](#) [BEYOND
RESEARCH](#) [ABOUT](#) [BLOG](#)

Let the mystery be

April 13, 2018

Iris DeMent sings, in her unique southern twang, a lovely song, (you can [find it here on youtube](#).) Its chorus goes:

Everybody's wonderin' what and where
they they all came from.
Everybody's worryin' 'bout where they're
gonna go when the whole thing's done.
But no one knows for certain and so it's all
the same to me.
I think I'll just let the mystery be.

I'm sure nothing I have to say about the deep questions of life is original. My recently deceased wife Jenifer was never predisposed to reflect like this. She truly lived in the present and withstood a progressive disease for 9 years by accepting each limitation as it came, sustained by friendships and love. But my curse, like many scientists, is to seek answers to questions that have no answers and writing my thoughts down seems to be part of the process. This post is a meditation on consciousness, the

The Dismal Science and the future of work

July 20, 2015

Let the mystery be

April 13, 2018

Can an artificial intelligence machine be conscious?

April 11, 2019

Can an artificial intelligence machine be conscious, part II?

July 12, 2019

Letter to my Grandchildren

March 1, 2020

Ridiculous Math Problems

mind/body problem and what science has said about these issues, and then touches at the end on the nature of the soul.

Consciousness is the thing that starts up in each of us when, as in the song, we 'come from' a place that no one knows and leaves us when 'the whole thing's done'. Many people have sought theories of consciousness. I recently read an op-ed piece in the New York Times that I especially liked. Instead of the antiseptic word 'consciousness', the author, Sean Kelley calls his piece "Waking up to the Gift of 'Aliveness'". The article is a commentary on the sentence "The goal of life, for Pascal, is not happiness, peace, or fulfillment, but aliveness" that he traces in some form to his teacher Hubert Dreyfus. He confesses that he knows no definition of aliveness but gives us two examples: looking at your lover's face when you have fallen in love; and lecturing to a class (he is a Professor) when your students are truly engaged and the classroom is buzzing. I take it that aliveness should be thought of as the most fully realized states of consciousness. While consciousness is the substrate of everything we do when we are alive in the mundane sense, the aliveness he is talking about is found in its most real moments, when all of life feels like it makes sense. He says aliveness should have the passion of Casanova without his inconstancy and the routine of Kant without his monotony. I'd like to think this is also the state of an enlightened Buddhist during meditation. And for me, I think this was how I have felt sailing with my spouse, when the physical, the mental and the emotional strands of life all wove together.

April 1, 2020

Reading Spinoza

April 19, 2020

Nothing is simple in the real world

July 8, 2020

The Astonishing Convergence of AI and the Human Brain

October 1, 2020

Ruminations on cosmology and time

March 1, 2021

The Shape of Rogue Waves

June 27, 2021

Draft of a *Blog Book*

August 16, 2022

Black hats and white hats

December 12, 2023

Consciousness, Robots, and DNA

January 3, 2024

STATEMENT

April 1, 2024

Letter to my Grandchildren-2

August 30, 2024

Science struggles with the idea of consciousness

Black hats and white hats-II

January 20, 2025

Now science has a real problem here: for a long time, even the word consciousness was taboo to practicing scientists. When I was a student, psychology had been overtaken by behaviorists and biology was being reduced to biochemistry. In this atmosphere, the mind/body problem had been left to philosophers (and a few quantum physicists -- see below). The first time I encountered the taboo breaking was when I read the neuroscientist John Eccles' 1977 book, joint with the philosopher Karl Popper, entitled *The Self and its Brain*. Both Popper and Eccles are believers in a *Three World* view of reality: (I) the objective physical world, (II) the inner world of conscious beings and (III) the world of ideas, that is objects of thought. Concerning the first two, they sought a detailed model of how in particular the physical brain interacts with conscious experience. Eccles developed their ideas further, e.g. in his 1990 paper *A unitary hypothesis of mind-brain interaction in the cerebral cortex*, Proc. R. Soc. London. His hypothesis is first that the cerebral cortex can be broken up into about 40 million columnar clusters, each made up of about 100 pyramidal cells which stretch from near the inner to the outer cortical surface, clusters that he calls *dendrons*. Secondly, each dendron 'interfaces' with a corresponding unit of conscious thought that he calls a *psychon* via an interaction allowed on the physical side by quantum uncertainty. Two figures from his Royal Society paper are reproduced in the sidebar. This is a breathtakingly bold and precise answer to the mind/body problem but one that

has not drawn many adherents,

More recently, scientists realized they could study *access consciousness*, that is the stuff that a person reports they are thinking, as opposed to consciousness as the ineffable, subjective sense of being alive, and then consciousness became something on which they could do experiments. Of course, they now exclude things like the reportedly heightened consciousness of Buddhists deep in meditation when all distracting thoughts of accessing the rest of the world are put aside. The goal of this research is to elucidate the *neural correlates of consciousness* with the aid of tools like fMRI, i.e. can one identify the precise neural states in which a person will report being conscious of something. It turns out that this is not as simple as one might hope: there are many sensations that cause measurable activity in primary sensory and other cortical areas that people are consciously unaware of and there are neat experimental ways of producing them such as binocular rivalry, masking, and attentional distraction. Even the order in which two sensations occur can be experienced consciously as the opposite of what actually happened. Strikingly, the conscious decision to do an action seems to occur *after* there is brain activity initiating the action. Moreover a certain amount even of reasoning can also be accomplished quite unconsciously by the brain. Freud would have told them that even strong emotions and actions resulting from these emotions often do not reach consciousness -- but his work was another taboo to scientists. The limitations of the self-awareness that consciousness provides were clearly summarized in Alex Rosenberg's

NY Times piece *Why you don't know your own mind*. His conclusion is "Our access to our own thoughts is just as indirect and fallible as our access to the thoughts of other people. We have no privileged access to our own minds".

What does make things conscious, according to many neuroscientists, is that the activity expressing some thought should spread over large parts of the brain, an idea known as the *global workplace theory* of consciousness. Almost the exact opposite of Eccles' theory, this proposes that activity over large parts of the cortex, often synchronized via ~ 40 Hz brainwaves (so called gamma oscillations) and where each part of the cortex contributes to the thought, is necessary and sufficient for the thought to be conscious. By every area, we mean the primary sensory areas can be involved but also the pre-frontal cortex and the so called association areas of parietal cortex perhaps have to be involved. I recommend Stanislaw Dehaene's book *Consciousness and the Brain* for a detailed description of this theory. Dehaene writes:

"Consciousness is like the spokesperson in a large institution. Vast organizations such as the FBI, with their thousands of employees, always possess considerably more knowledge than any single individual can ever grasp. ... As a large-scale institution with a staff of a hundred billion neurons, the brain must rely on a similar briefing mechanism. The function of consciousness may be to simplify perception by drafting a summary of the

current environment before voicing it out loud, in a coherent manner, to all other areas involved in memory, decision, and action.(op.cit. p.99-100)

How about animals?

Access consciousness cannot be used to interrogate animals without speech and, in any case, it hardly captures our experience of aliveness. Just as we reject solipsism and feel sure our fellow humans have the same experience of consciousness as we have, we observe, for example, our pets and see their awareness of the world, see their emotions and their awareness of our emotions. Most people readily conclude that they share our sense of aliveness too, and refuse to believe that what goes on in their brains is a mere mechanistic biochemical process lacking the spark of subjective awareness. Moreover, speech is not all that reliable an indicator that we are in touch with another sentient being anyway: computers have occasionally been able to pass the Turing test and fool observers into thinking a real person is talking to them over a phone. So speech alone is an unreliable token of real consciousness.

What can a scientist use for assessing consciousness in mute creatures? The main experimental tool in testing monkeys has been to train them to respond to a stimulus in different ways, e.g. by pressing various buttons, assuming that producing such a response means that the stimulus has activated something we can call their consciousness. In this way, a whole body of research has confirmed that consciousness in monkeys

follows patterns similar to that in humans. For example, some stimuli do not reach consciousness and when they do, large parts of the monkey's neocortex show activity, often synchronized by gamma waves.

All mammals have virtually identical brains, differing only in the size of its constituent parts. Thus human brains are distinguished by having a greatly enlarged pre-frontal cortex that appears to parallel the greatly increased activity in which the mind is making complex plans. On the other hand, non-mammalian vertebrates have brains which are all fairly similar to each other and similar to the mammalian brain *if you remove the neocortex*.

The neocortex has a unique 6-layered structure not found in non-mammals although most recent thinking is that its parts are present, just not assembled and wired with pyramidal cells (as in Eccles' theory) as they are in mammals. These parts, called the pallium in birds and just the cerebrum in all classes, the 3-layered paleocortex, especially the hippocampus, as well as the thalamus (sometimes considered as a seventh layer of neocortex) are found in other vertebrates. So how far down the evolutionary tree can one make a convincing case for consciousness? Many people are convinced that birds, especially parrots and crows, are conscious beings.

In the previous section, we reviewed the global workplace theory that puts the neural trace of consciousness squarely in the neocortex. This leads you to conjecture that consciousness may exist in all mammals and only in them. But another theory goes back to Penfield's

experiments operating on patients with epilepsy. To locate the exact cortical area whose excision would cure the epilepsy, he operated with local anesthesia and interrogated his awake patients while stimulating their exposed cortices on the operating table. In this way, he almost always found some area where the trigger for the epilepsy was located. But one form of epilepsy, *absence epilepsy* in which the patient briefly loses consciousness without any other symptoms, did not correspond to any unusual cortical electrical activity. This led him to propose that consciousness is related not to the neocortex but instead to activity in the *midbrain*. This theory has been extended by Bjorn Merker (see his paper *Consciousness without a cerebral cortex*, Behavioral and Brain Sciences, 2007) who filmed and worked notably with hydranencephalic children, children born with no neocortex (though the paleocortex and thalamus are usually preserved). His claim is that when given full loving custodial care, and within the limits imposed by their many weaknesses, they exhibit behavior much like normal children. For me, however, a really stunning piece of evidence for this theory was the 'Turing test' of this issue was that carried out by Jaak Panksepp (see Panksepp's commentary to Merker, op. cit., pp.102-103). He surgically removed the neocortex in 16 baby rats, paired them with normal rats and asked 16 of his students to each watch one such pair play and guess which rat was intact, and which lacked their neocortex. Only 25% of normals were correctly identified, while the decorticates were judged to be normals 75% of the time! It

seems the major role of the neocortex was to make rats that possessed one more cautious, leaving the decorticates more playful. The bottom line is that, if you subscribe to the midbrain location hypothesis, one should ascribe some form of consciousness to all vertebrates.

On the positive side, research into bird behavior and biology has lead many people to believe that birds, especially crows, parrots and owls, are everybit as intelligent and responsive as, e.g. dogs and cats. A wonderful review is the book by Jennifer Ackerman, *The Genius of Birds*. So far we haven't gone beyond vertebrates and touched on octopuses. I knew the famous neuroscientist Jerry Lettvin who worked with octopuses and (personal communication) was convinced that they were conscious beings and loved playing practical jokes on him. This has been confirmed by many observers. It seems they enjoy immensely playing with human toys -- see sidebar. A beautiful book, *The Soul of an Octopus: A Surprising Exploration into the Wonder of Consciousness* by Sy Montgomery develops this thesis drawing on extensive personal interactions (or should I say 'relationships') with octopuses. See also [this wonderful lecture by Montgomery on her experiences](#). Personally, I find it quite convincing that indeed birds and octopuses have consciousness. To put this in context, octopuses have about 500 million neurons, in the range 200-2000 million neurons that includes rats, cats, dogs, crows and owls (humans have 100 *billion*, though only some 20 billion in neocortex). I believe it is only reasonable to ascribe some form of

consciousness to all these animals, a slippery slope to be sure but the evidence, based on relationships with people, is strong.

Physicist's quandary and the 'Bohr bubble'

As mentioned above, physics has also grappled with the concept of consciousness. To explain this, we need a few ideas from quantum mechanics. To model the subatomic world, quantum mechanics uses wave functions. These are typically small sets of complex-valued functions of the spatial coordinates of all the particles present $\psi_\alpha(\vec{x}_1, \vec{x}_2, \dots)$ called Schrödinger wave functions. The details don't matter. What does matter is how the wave functions relate to the world: are they ontological, describing objective material reality or are they epistemic, describing an observer's knowledge of the world? The problem is that they are both! They are ontological in the sense that it has been shown, convincingly to essentially all physicists, that there can be "no hidden variables", meaning any more detailed descriptions of the state of the world than the wave functions ψ . They are "all that there is" when looking at an atom. Yet they are epistemic in that if, in a lab, some aspect of a subatomic event has been amplified and then is observed by a human, then that human knows something new and he/she must reset the wave function if they wish to best predict future events. This quandary has disturbed Heisenberg, von Neumann, Bohr, Einstein, Feynman -- all the great physicists. An excellent discussion can be found Eugene Wigner's essay *Remarks on the Mind/Body Question* (reprinted in his book *Symmetries*

and Reflections). There has been a trend in the last few decades to say that something called *decoherence* has solved the problem, but all this does is show how a subatomic level superposition $\psi_1 + \psi_2$ of 2 states reduces on the macroscopic level to a choice between 2 throws of a die (technically speaking, choosing one of the eigenvectors of a 2×2 density matrix), each having some probability. The outcome of throwing the die still forces a sudden reset of the quantum state.

The meaning of ψ is part of the broader question of reconciling an inherently indeterminate description of subatomic events with the determinate classical description of macroscopic events. Physicists almost religiously resist the conclusion that human consciousness must enter into the reconciliation of these two descriptions. But Wigner and many others are resigned to this. Wigner writes "The preceding argument for the difference in the roles of inanimate observation tools and observers with a consciousness ... is entirely cogent so long as one accepts the tenets of orthodox quantum mechanics in all their consequences." A key point, however, is what happens if there are two physicists A and B, A making measurement A' and B, later in another room, making a second measurement B'. The probabilities of the outcome B' will be altered by the outcome of A', so the resulting ψ must reflect both measurements. This makes the epistemic viewpoint reflect the joint knowledge of both physicists. Pushing these ideas further, one is led to believe that our whole civilization is creating a bubble in space-time in which ψ is forced to reflect all the measurements all of us

have done, the deterministic realities of our lives. As this is all a consequence of what physicists call the "Bohr" or "Copenhagen" interpretation", I call this our Bohr bubble in which, weirdly enough, our consciousnesses have altered the objective world..

Consciousness defines what we call *now*, science doesn't

From my point of view, however, I think all scientists are missing the essential nature of consciousness. Sure we are conscious of what our eyes see and our ears hear, sure we are conscious of moving our body and making plans to do stuff and sure we can even fill our consciousness with the imaginary world of a novel or the proof of a theorem. But I think all this misses what makes consciousness absolutely different from anything material: consciousness creates for us a present moment and it does this continuously moment after moment. I hear you say, isn't this just a part of the physical nature of the world? After all, Newton said "Absolute, true, and mathematical time, of itself, and from its own nature, flows equably without relation to anything external". Actually, Einstein's special relativity says this just isn't true: time and space have to be joined together into a single 4-dimensional space-time world. This is because there is no physical meaning to two events being simultaneous when they occur in different places because there is no consistent ways to synchronize two clocks carried e.g. on different rockets. Therefore there is no universal time and you cannot measure time independently from measuring space. Not just that but science, essentially by definition, only

studies correlations between events that can be reproduced exactly enough to show something is repeating itself, that here is a law of nature valid at least throughout some region of space-time. Thus it refuses to deal with any special instant that someone might call "now". The word "now" only enters our vocabulary through our conscious experience. For us, an experience is *never* reproducible (though we often try to make it so). As the saying has it, "you only go round once".

Sure, physics studies unique events such as the explosion of the Crab nebula seen on Earth in 1054 CE, but this is a fact of history, not a scientific law. The science of astrophysics explains this explosion by equations which are then applicable to infinitely many stars and in this way removes the historical uniqueness of that supernova. It's interesting to recall a famous debate between Einstein and the philosopher Henri Bergson (thanks to my son Peter for telling me about this). They met in 1922 arguing about the nature of time. For Bergson, the important notion of time was not that of clocks but that of people's immediate conscious experience (*Les Données Immédiates de la Conscience* was the title of his dissertation). When you focus on the subjective time of a person, it is indeed not bound up with his spatial location. I find his ideas hard to follow but I think the key one is that time is heterogeneous, not homogeneous. Each instant for a conscious being is a thing in itself and their totality cannot be counted, like a flock of sheep. Time, he says is a temporal heterogeneity, in which "several conscious states are organized into a whole, permeate one another, [and] gradually gain a richer

content" (Stanford Encyclopedia of Philosophy). In contrast, Einstein would say that a person's lifetime is a curve in space-time, time-like meaning the person moves more slowly than light, bounded by the space-time points representing his birth and his death, and along which integrating the Lorentz metric computes each person's subjective time. You can see these guys are not going to reach a consensus. Apparently Bergson's denial of Einstein's theory of time was the reason his Nobel Prize was awarded instead for his work on the photo-electric effect.

I find philosophical writings like Bergson's awfully hard to follow. But one thing seems totally clear to me and this is the central point of this post. Science only studies correlations between events that can be reproduced, independent of time. On the other hand, the conscious life of a person is at all times structured into past, present and future. The word "present" has no meaning in physics. Physics typically sets up situations in an abstract time and space that can be realized in many space-time points of the real world. I want to argue that it is this experience of a present instant, the *now* that is always changing yet is always our one and only unique present, the one that each of us owns, that that is the real core of what we call consciousness. You see this explains the Buddhist meditator: his mind may be empty of worldly distractions and his cortex may have no sensory, motor or memory activity but he still lives fully his present moment.

Spirit? Soul?

I am aware that when you talk of these things, you place yourself in the firing line of reductionists and atheists, irritating people like Richard Dawkins. Dawkins is one of the most vocal and most scornful denouncers of so called wishy-washy ideas, saying for example:

"The meme for blind faith secures its own perpetuation by the simple unconscious expedient of discouraging rational inquiry."

(http://www.age-of-the-sage.org/quotations/quotes/richard_dawkins_god.html).

Nonetheless, forging on, I think that once we acknowledge the simple subjective truth that we are all living in the *now*, it means that you are acknowledging that mind cannot be reduced to body and it gives what I consider to be a perfectly rational need to take seriously things like spirit and soul. Once you leave aside the *content* of mind, the daily grind and its pleasures and sorrows, there seems to be something more, a pure awareness still active in the mind. I feel we need to reach into our language and tradition to find its proper name. For me, 'spirit' seems about right. I am drawn to describe the mind/body dualism as being the result of the spirit-mind falling in love with the world-body, the spirit choosing to embody itself for the sheer joy of bathing in the love (the hate too?), all the passions that life delivers, of being bound to the world by an infinitude of ties.

Does this spirit, in and of itself, have any attributes? For example, does your spirit have specific skills or disabilities, likes or dislikes, or is it attached to a specific places as an alive person does? These all seem to be accidents of an incarnation and I think it's not hard to imagine the life you would have had if you

were brighter/dumber, more beautiful/more ugly, or born in some other country. If you can imagine many different lives like these, then presumably your spirit is not indissolubly tied to these specifics either. There is surely no reason to give your spirit a gender either. With an effort of empathy, an activation of suitable mirror neurons, I don't think it is all that hard to imagine something of what it would have been like had we been born in the opposite sex. A much harder question is what relationship should spirit have to time. Since experiencing time is how I have argued that we know spirit is there, is this experience the result of the embrace of a body by the spirit or did spirit bring this to the union? Here we touch the question of the existence of an 'after-life'. Why don't we ask equally often about the existence of a 'before-life'? We are surely back now at DeMent's "Let the Mystery Be". We have no rational way of thinking about spirit in an unembodied state and therein lies what for me is the ultimate mystery. People cavorting in heaven in white gowns certainly seems ludicrous as it would seem that all the trappings of individual identity are merely corollaries of embodiment and not inherent characteristics of spirit.

Computer people are fond of acronyms. My personal creation is to describe human beings as SOUL machines, using an abbreviation of the nominal phrase "Single Opportunity for Use Learning" machines. Talking of computer people, their big question is whether consciousnesses and spirit will ever be born into a robot. If enough smarts (and guile?) are built into the 21st centuries silicon "monster", will it (he/she?) become conscious, as in all

those recent movies? In Hollywood, Frankenstein's consummately ugly creation has been replaced by shiny curvaceous seductresses! But this blog has gotten too long already.

This blog has been translated into Slovenian by Andrijana Savicevic: [click here for the Slovenian translation](#) . Thank you Andrijana. And, more recently, it has been translated into Ukrainian by Sandi Wolfe: [Click here for the Ukrainian translation](#). Thank you Sandi.

David Mumford's content on this site is available under a [Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License](#). [Click here for sitemap](#)
