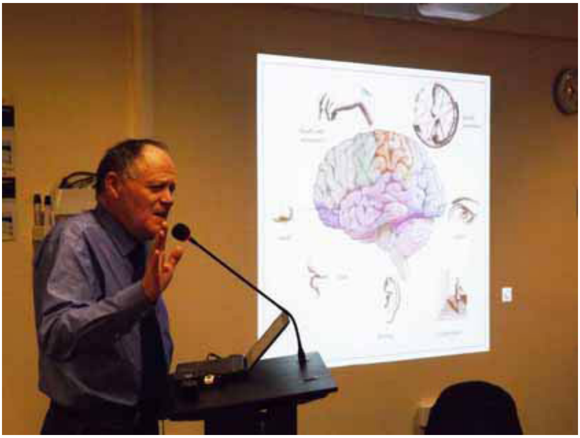
London Newman Lecture 2015
Neuroscience and the Soul

By James Le Fanu



Neuroscience is so vast a subject that reflections on its contribution to our understanding of the metaphysical must inevitably tend to the overly simplistic. As for the soul, while those attending the London Newman Lecture might be expected to believe in its existence, there is likely to be a considerable difference of opinion as to precisely what that belief entails. Nonetheless my foolhardiness in addressing this subject is justified by the fact that in the recent past the details of the findings of neuroscience have quite unexpectedly illuminated in several interesting ways the common, if not strictly theological, perception: that the soul is that unique character or personality that we know ourselves to be.

The central philosophical question I will be addressing can be simply put. The brain, by common consent, is qualitatively different from any other organ of the body. Standard textbooks of cardiology or respiratory medicine describe in almost exhaustive detail the working of the heart and lungs, but the fundamental question of what the brain does remains quite unresolved – how do those neuronal circuits give rise to those non-material properties of our mind: our thoughts and imaginings, joys and sorrows and the sense of self.

The prevailing view of course is that those thoughts and imaginings and the sense of self must ultimately be explicable in terms of the physical properties of the brain as eloquently set out by Colin Blakemore, formerly Professor of Theology at Oxford University. He said: "The human brain is a machine which alone accounts for all our actions, our most private thoughts, our beliefs. It creates the sense of self. We may feel ourselves to be in control of our actions, but that feeling is itself the product of our brain whose machinery has been designed by natural selection".

The insuperable intellectual difficulty with this reductionist interpretation of a relationship between brain and mind is that it denies by necessity the reality of the two central and most important, aspects of our experience – our sense of personal identity and free will. Rather, by the materialist account, as Blakemore makes clear, the sense of self and free will can only be an illusion generated by the brain to create the impression that someone is in charge. We are, in short, the stooges of our brain.

Intuitively we know this cannot be so but it is also of considerable interest to understand why this pervasive and influential view should be in error. Here context is all, so before turning to the recent findings of neuroscience it is necessary to consider, if briefly and schematically, how the scientific understanding of the relationship between brain and mind has evolved over the past one hundred years.

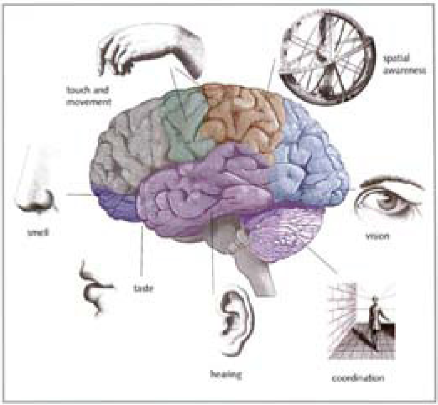


Fig 1 – The specialisation of the parts of the brain to fulfil its different functions includes (as illustrated here) the sensory perception of vision, smell, hearing and taste. The frontal lobes are dedicated to "higher" attributes of reason and imagination, and large tracts of the left hemisphere to language.

The three-pound entity

The human brain is the most capacious entity in the universe, both transcending *time* – recalling the past, experiencing the present, anticipating the future – and encompassing every magnitude of *space* from the vastness of the cosmos to the near intifessimal smallness of a single atom. The brain also poses the greatest conundrum within that universe – how those three pounds of protoplasmic stuff could give rise to the distinctive character and personality ,'the soul' of each one of us, both the billions with whom we currently share this planet and all who have gone before.

And more, how it is that, moment by moment, that same protoplasmic stuff perceives the world out there in all its exquisite detail, stores its experiences as memories to be recalled decades later and makes sense of the world through the powers of reason. How can the material brain be the causal basis of so vast a range of mental life? The obvious answer is that it does not and cannot and for more than 2,500 years that dissonance between the material brain and the properties of the mind was the most persuasive evidence for there being a dual nature to reality – a material and non-material domain.

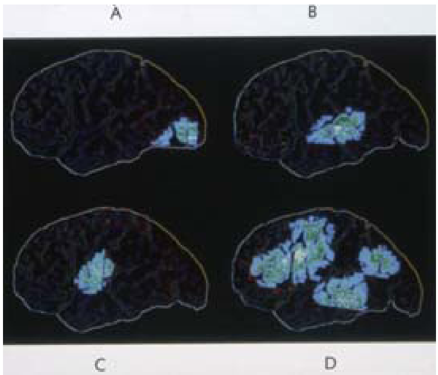


Fig 2 – The simplest of intellectual tasks such as reading or listening to the word "chair" generates widespread electrical activity involving millions of neurons in the visual and auditory cortex. Associating the noun "chair" with verb "sit" activates further vast tracts of the brain.

The first modern philosopher Descartes made this distinction with great clarity, pointing out how material objects such as the brain occupy space and are objectively knowable, while the non-material elements – our thoughts and imaginings – do not and are only knowable to their possessor. Now, brain and mind must obviously be linked as, self-evidently, injury to the brain impairs the workings of the mind; yet it has always been part of the commonsense of mankind that body and soul, brain and mind are two different things.

This dualist interpretation could scarcely survive the ascendancy of science in the 19th century. Darwin famously in *The Descent of Man* incorporated humans into the evolutionary framework by denying the exceptionality of the human mind, arguing that its attributes were different only in degree but not in kind from those of our primate cousins. His great supporter Thomas Huxley was less convinced, suggesting that the mystery of how the brain gives rise to the mind is just as inexplicable as the appearance of the Djinn when Aladdin rubbed his lamp. Still, there seemed no alternative if science were to proceed to suppose a direct causal relationship between brain and mind; indeed, soon afterwards two important discoveries suggested the human brain was not nearly as inscrutable as it might appear.

First, the French neurologist Pierre Paul Broca noted at the autopsy of one of his patients, known as Tan – because that was the only sound he had been able to utter since his stroke 30 years previously, that the patient had suffered a discrete area of damage in the posterior part of his left frontal lobe that he inferred to be the speech centre of the brain. Soon after the German physician Karl Wernike described a comparable defect in a patient unable to comprehend speech. Over the next sixty years similar natural experiments revealed the cerebral hemispheres to be a chequerboard of specialised functions with which we are now readily familiar (see Fig 1).

This cartographic map of the brain can however be misleading on two counts. Firstly, while it is possible to allocate specialised functions to these discrete parts of the brain, large areas of the frontal and parietal cortex remain unaccounted for. From this one might infer they have some integrative role related to the higher functioning of the mind. Secondly, this phenomenon of localisation, fascinating as it is, offers no insight as to how it translates into subjective experience: how does the electrochemical activity of the millions of neurons of say, the visual or auditory cortex, give rise to such qualitatively diverse subjective experiences as watching a sunset or listening to a Bach cantata?

Brain or computer?

This brings us to the second stage of the unravelling of the relationship between brain and mind, from 1950 onwards, when the synthesis of four observations suggested a much more sophisticated metaphor for the brain, not as a map but as an information processing device – or computer.

First, the British neurophysiologist Charles Sherrington demonstrated how the individual neurons in the brain have two modes – where they can *excite or inhibit* the electrical activity of other neurons in close proximity. There is a clear parallel here with the brilliant Alan Turing's conception of a universal calculating machine that could in principle carry out any mathematical task using a binary code of just two symbols – 1 and 0. Further, the rich, dense circuitry of the brain is readily comparable to a microprocessing 'chip', and there is an obvious analogy between the famous distinction between the relative contribution of 'nature' and 'nurture' to the workings of the human mind and with the hardware and software of a computer.

This computer metaphor has certainly proved very fertile. The neuronal connections of the brain, its hardware, come hard-wired at birth with specific modules for language, music, mathematics, the ability to recognise faces and much else besides. As for the software, the phenomenal neuroplasticity of the brain ensures the young brain voraciously and remorselessly programmes itself, integrating into its workings the culture, emotions, sights and sounds that it encounters. The similarities are immensely compelling, but while the computer's power to crunch numbers is one thing, the capacity of the human mind to hold a conversation or feel happy or sad is another. And here, when put to the test, the more closely the comparison is pursued the more astonishing and un-computer-like the brain appears to be.

This brings us to the third phase: the neuroscientific revolution of the recent past made possible by the novel and astonishing scanning technology that allowed scientists for the first time to observe the brain in action *from the inside*, thinking, perceiving and reflecting on the world out there. It all began just over 25 years ago in 1988 with a scientific paper in the prestigious journal *Nature* by two American scientists Marcus Raichle and Michael Posner: *Positron Emission Tomography (PET) Studies of the Cortical Activity of Single Word Processing*.

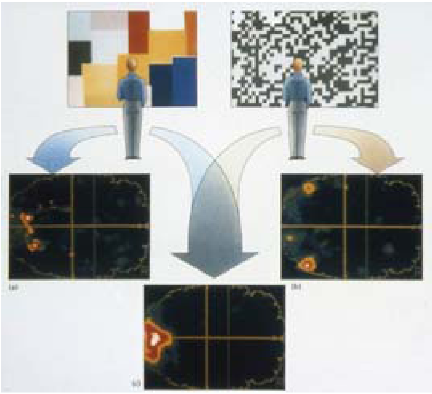


Fig 3 – The PET scans of subjects when viewing a colourful Mondrian painting (a) pinpoint the visual map for the interpretation of colour. By contrast the brain scans of subjects viewing a series of moving squares (b) activate the visual area concerned with movement.

The principle behind PET scanning technology lies in the fact that when, for example, a person is speaking the neurons of Broca's language centre in the left frontal cortex massively increase their demand for oxygenated blood reflected in an increase in blood flow. This can be detected by labelling the oxygen molecules with a radio isotope that is detected by the PET scanner and converted by ingenious mathematical algorithms into multi-coloured images.

There can, for example, be no simpler task than that investigated in this first paper on 'single word processing' – scanning a subject's brain when reading, hearing or repeating a single word such as 'chair', a trivial task indeed which nonetheless was shown to generate a blizzard of electrical activity across the relevant visual cortex and the language centres of the brain. When that task was made slightly more complicated by asking the subject to associate the noun 'chair' with a verb 'sit' there were activated, in addition, vast tracts of the frontal and parietal lobes (see Fig 2). Who could have supposed that this apparently simplest of tasks appears to involve the brain virtually in its entirety? What, one might reasonably enquire, must be going on in the brain during even the most elementary of conversations?

It would be impossible to summarise the findings of the flood of scientific papers generated by this novel technology but the most revealing – certainly in radically changing our understanding of the workings of the human mind – fall into the four main categories already alluded to. The first is perception: how the brain, through the senses, perceives the world out there in the most exquisite detail. Next comes memory, the lynchpin of the human mind, holding the past and present in a permanent embrace. Then we have free will, or mental causation: how it might be that our non-material thoughts can nonetheless affect the physical workings of the brain so as to compel us to take one action rather than another. And, finally, there is the power of reason as revealed by its mediation through the faculty of language.

Fragmented images

We start with **perception** and one sense in particular, vision. The conventional and very persuasive view would be that the image of the world 'out there' is captured by the retina at the back of the eye and imposed, like a photographic plate, on the visual cortex. But perhaps the most dramatic of all those recent findings of neuroscience is that, on the contrary, the visual cortex fragments that image like an exploding firework. This was elegantly illustrated by Professor Semir Zeki of the University of London in a classic experiment that involved scanning the brains of volunteers looking first at a multi-coloured Mondrian painting and then at a screen filled with moving black and white squares to reveal two distinct 'hot spots' involved in colour and movement respectively. (Fig 3) Further investigation has revealed the brain perceives the world 'out there' by fragmenting it into thirty separate specialised functions scattered throughout the cortex.

Clearly, those fragmented functions must be reintegrated back into that unified stream of perception by means of which, moment by moment, we perceive the world out there – but how? David Hubel, Nobel prizewinner for his investigation of vision, clarifies: "This abiding tendency for form, colour and movement to be handled by separate structures in the brain immediately raises the question about how all the information is finally assembled, say for perceiving a bouncing red ball. It obviously must be assembled, but where and how **we have no idea**".

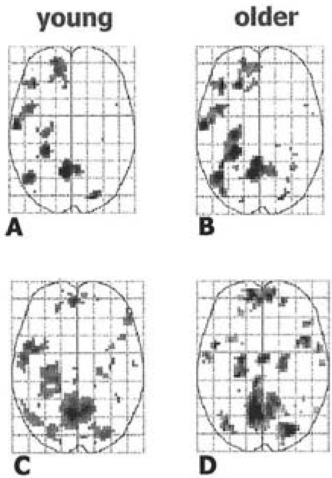


Fig 4 – Overlapping memories. A simple memory task lights up large tracts of the brain in both young and old, but there is a surprising degree of overlap in the retrieval of general knowledge (top) and a very different recall of autobiographical events (bottom).

Next we turn to memory, the lynchpin of the human mind with its three distinct components of memorising, storage and retrieval; this comes in the three forms of short, medium and long term and as two distinct types – declarative (the memorising of facts) and autobiographical (the memory for past events). The conventional view, invoking the computer analogy, would be that memories are stored away each in its own neuronal circuit and are available for recall at a later date. But that emphatically is not the case: this was revealed by a study by Professor Eleanor Maguire of the Institute of Neurology, who investigated these two very distinct types of memory – the factual and autobiographical (see Fig 4). Two groups of subject, young and old were asked first to recall the knowledge of certain facts (that, for example, the sun is 92 million miles from the earth) and then autobiographical (such as their starring in the school Nativity play). (Fig 3). Here we note first that large overlapping areas of the brain are involved in both tasks suggesting that these two very distinctive forms of knowledge must share many of the same neuronal circuits. Then it would appear, however, that as we get older the existing forms of memory are memorised and recalled in different parts of the brain. This is profoundly puzzling.

The most striking feature of memory is its fixity – the ability to recall instantly and effortlessly facts and events from forty years or more ago. But the impression conveyed by these scanning techniques is precisely the reverse: even the simplest memory involves overlapping tracts of the brain, the localisations of which are surprisingly fluid, shifting from one part of the brain to another.

Next we turn to 'free will', the powerful impression that our freely-chosen thoughts can influence our actions; as noted, Professor Colin Blakemore insisted this must be an illusion generated by the brain. This would seem to contradict not just our everyday experience but specifically the rationale of psychotherapy, which is predicated on the assumption that mental states can be influenced for the better by thinking and reflecting on them. Psychiatrist Professor Jeffrey Schwartz of the University of California has demonstrated this in a study of patients with obsessive compulsive disorder in whom brain imaging studies identified increased activity in the part of the frontal lobe known as the caudate nucleus. Following ten weeks of cognitive therapy their symptoms of OCD were much improved, paralleled by 'normalisation' of their abnormal brain activity.

As Professor Schwartz himself has put it: "We have demonstrated the sort of changes that psychiatrists might see with powerful mind-altering drugs, but in patients who had changed the way they thought about their thoughts." So, *contra* Professor Blakemore, non-material thoughts can, it would appear, influence the physical structure of the brain.

Finally we turn to the human mind's power of reason by which it makes sense of the world which, as we all know, is inextricably linked to the faculty of language that allows us to 'think' by assigning words to objects and ideas and then applying grammatical rules to the arrangement of those words. Darwin in his influential *Descent of Man* insisted that human language evolved from the grunts and groans by which our primate cousins still communicate with each other. But language cannot just be an evolved module of the brain when, as will be recalled, the simplest of linguistic tasks such as associating the word 'chair' with the word 'sit' involves activation of the brain virtually in its entirety. Further investigations would reveal that, as with vision, the brain fragments words into their constituent parts with specialised areas involved in naming letters or words – posing precisely the same problem addressed by David Hubel as to how they might be reintegrated back together into a flow of conversation.

The five cardinal mysteries of the mind

So what to make of this? Neurobiologist Robert Doty, reflecting on the findings of these brain imaging studies, has drawn attention to what he describes as "The five cardinal mysteries of the mind". They are:

- The mystery of subjective experience – how the monotonous electrochemical biology of the brain should give rise to the infinite variety of the subjective experiences that fill our lives, from the scent of a rose to the cadences of a Bach cantata.

- The mystery of free will – how our nonmaterial thoughts and intentions impel us to take one course of action rather than another.
- The mystery of the richness and accessibility of memory. Here Robert Doty observes: "This facility to sort with alacrity amongst the items of a lifetime, pursue in milliseconds obscure, half-forgotten episodes and their cascading associations defies credible clarification".
- The mystery of human reason and imagination by which, through the faculty of language, we make sense of the world in which we live.
- The mystery of the sense of self – with its distinctive character and personality that may change and mature over time but remains in essence the same.

And now here is the crunch: these may be mysteries to science but they are certainly not to ourselves. Indeed, there is nothing we can be more sure of than our subjective experiences, memories, free will, powers of reason and sense of self. That unbridgeable gap between the limitations of those objective investigations of the workings of the brain and our personal knowledge of our everyday experiences eliminates any possibility of there being an adequate scientific explanation for the relationship between brain and mind.

And as that is the case, the materialist assumption that the sense of self and free will can only be an illusion generated by the brain ceases to be tenable – and so by inference the non-material soul must be for real.

This London Newman Lecture 2015 was delivered at Heythrop College on March 12th

Dr James Le Fanu is a general practitioner and a journalist, with a regular column in the *Daily Telegraph*. He has also written books, including *Why Us? How Science Rediscovered the Mystery of Ourselves*, published by HarperPress.