Brains and minds: a brief history of neuromythology

R C Tallis

Introducing neuromythology

At the heart of neuromythology is the belief that the brain is the seat of the soul; more specifically, that neural activity explains consciousness. I believe this is wrong, not only regarding the nature of human consciousness but regarding the place of human beings in nature. Neuromythology is consistent with what the philosopher Daniel Dennett has called ‘the prevailing wisdom’. This is the belief that:

There is only one sort of stuff, namely matter – the physical stuff of physics, chemistry and physiology – and the mind is somehow nothing but a physical phenomenon. In short, the mind is the brain. We can (in principle) account for every mental phenomenon using the same physical principles, laws and raw materials that suffice to explain radioactivity, continental drift, photosynthesis, reproduction, nutrition and growth.

It is consistent with this outlook that neuroscientists, and philosophers of mind impressed by neuroscience, have been inclined to believe that as we learn more about what happens in the brain, we are getting closer to understanding how the mind works; more fundamentally, what mind is and how it comes to exist. This is misleading about the scope and limits of science and this ‘scientism’ may give science a bad name.

Against neuromythology, I will argue that not only has neuroscience failed to cast light on how there is such a thing as the mind but also that it is unlikely ever to do so. In arguing this, I do not intend to diminish the spectacular and important advances that have been made in understanding the workings of the brain, rather to deny that they give us any insight into the nature (or the basis) of consciousness, the mind or, if you will, the soul of man.

The modern belief that the activity of the brain provides an adequate explanation of human consciousness has several connected elements:

- that mental phenomena are identical with neural activity (or patterns of neural activity) taking place in certain parts of the brain
- that in the case of perception, this activity is caused by energy impinging on the brain (the causal theory of perception)
- that the brain, in this regard, is like a computer (mind is the information processing activity of the brain)
- that mind/consciousness can be understood in terms of the evolutionary processes that gave rise to the brain.

Stephen Pinker’s statement that ‘the mind is a system of organs of computation designed by natural selection to solve the problems faced by our evolutionary ancestors’

gathers together many of these strands.

The brain as the seat of the soul

The history of neuromythology begins with a fatal mistake: the assumption that there is – indeed must be – an organ in the body wherein the soul or mind or consciousness is located. Although the theory antedated him, Hippocrates gave it its most striking expression. In On the sacred disease (epilepsy) he declared that:

Men ought to know that from the brain, and from the brain only, arise our pleasures, joys, laughter and jests, as well as our sorrows, pains, grief and tears. Through it, in particular, we think, see, hear, and distinguish the ugly from the beautiful, the bad from the good, the pleasant from the unpleasant ...

The brain is not only a necessary but also a sufficient condition of conscious experiences.

Powerful support for the central role of the brain in consciousness comes from the many ordinary observations that indicate that the condition the brain is in and the condition the mind is in are closely correlated. A bang on the head, with damage to the brain, may remove vision, impair memory or alter personality, all of which suggest that vision, memory, personality – everything from the most primitive buzz of sensation to the most exquisitely constructed sense of self – depend crucially on the functioning of the brain. For neuromythologists it follows that the mind or soul is housed in the brain.

The location of the soul within the brain

If we accept the notion that the mind is housed in the brain, or the soul is seated there, it seems reasonable to wonder whereabouts in the brain it is to be found. There was a very long argument as to whether the soul was in the parenchyma or the ventricles. Both possibilities were supported by ingenious arguments.

Over the centuries, ventricular and parenchymal theories
underwent considerable elaboration. According to one popular version, different ventricles housed different faculties of the soul: the anterior ventricle was the seat of fantasy or imagination; the middle ventricle the seat of reason; and the posterior ventricle that of memory. Eventually the notion that the parenchymatous tissue was the seat of the soul gained the upper hand but as recently as 1796, Soemmering (discoverer of the substantia nigra) argued – against Swedenborg – that the ventricular fluid was the repository of the soul.

By then, however, the parenchymal theory was unassailable and the question was where exactly within the parenchyma the soul was located. Thomas Willis was the first to suggest the cerebral cortex, a location favoured by contemporary neurophilosophers.

As speculative neurophilosophy, which drew on sources as disparate as religious doctrine, clinical observation, and rampant rationalising guesswork, gave way to what we would recognise as neuroscience, the question of the location of the soul became rather more complex. Perhaps (as the defeated ventricular theorists had believed) it was located in several places rather than one; or perhaps it was diffused over the entire brain, or over part of the brain such as the cerebral cortex. The decisive intervention was that of Gall, the co-founder of phrenology, who promulgated the following principles:

- the brain (especially the cortex) is the organ of the mind
- the brain is a composite of parts, each of which serves a distinct, task-specific ‘faculty’
- the size of the different parts of the brain, as assessed chiefly through the examination of the cranium, is an index of the relative strengths of the different faculties being served.

The third principle has dominated and damaged the phrenologist’s posthumous reputation, but the first two principles – the pre-eminence of the cortex in mental function and the localisation of different mental faculties within the cortex – have made an enduring contribution to the framework of neuroscientific research. Gall and Spurzheim were the first since Willis to identify the cortex as the basis of higher mental function. Moreover, the second principle addresses a serious philosophical problem.

Historians have linked phrenology with the need to deal with the problems arising out of John Locke’s theory of knowledge, promulgated in his enormously influential Essay concerning human understanding. Locke repudiated the notion of innate ideas and asserted that all knowledge came from the senses. The mind at birth was a ‘tabula rasa’ – a clean slate or a blank sheet – and was effectively constructed out of experiences organised only according to their associations. But if the mind was a blank sheet at birth, and built up out of experiences, how did it manage to avoid being just a heap of impressions, a slop of accumulated experiences and their echoes in memory, not too different from delirium – more of a Jackson Pollock than a Hughlings Jackson. There needed to be an innate material basis for the organisation of the material of which the mind was composed. Hence Gall and Spurzheim’s separate mental faculties associated with discrete organs in the brain.

With the advent of more sophisticated physiological experimentation, and the precise observation of both clinical and pathological aspects of neurological damage reported by authors such as Broca and Hughlings Jackson, the doctrine of localisation – in particular the localisation of functions within the cortex – became irresistible. The localisation of higher mental functions in the brain had already been dramatically suggested by the 19th century’s most famous neurological patient, Mr Phineas Gage, a railway worker who had an unfortunate encounter with a steel rod. This event, as the result of which he lost a lump of his frontal lobes, changed him from a purposeful, industrious worker, even tempered and impeccably mannered, into an evil tempered drunken drifter.

With the advent of modern methods of stimulating and recording from the central nervous system, of delineating its multifarious internal anatomical and physiological connections, and of imaging the living brain using a variety of modalities, we are now truly in a neo-phrenological era in which it seems as if every discernible function has its own piece of circuitry.

This trend towards localisationism has been driven by the conceptual and empirical advances set out in Table 1.

Today’s neo-phrenology is, of course, a long way on from the phrenology of Gall and Spurzheim. Not only is there now increasing emphasis on the plasticity of the brain, soft-wired modules and logic circuits rather than discrete anatomical sites, but also the functions into which the soul is fractionated tend to be things like object localisation, edge detection and encoding of episodic memory rather than the sense of justice or amatory propensity. The fundamental framework established by Gall and his later-19th-century successors, however, is the same.

**The problems of neuromythology**

Everyday life and neuroscientific observations all point to the inescapable conclusion that consciousness is due to

---

**Table 1: Conceptual/empirical advances supporting the notion of localisation.**

- the electrochemical nature of neuronal activity is recognised
- neurons are seen to be both discrete and interacting, through synapses, to form circuits
- the quantitative correlation between neuronal activity and sensory experience and motor outputs is demonstrated
- the preferential and modulated specificities of neuronal circuits – with filtering and amplification – is observed
- the notion of discrete modules evolves from anatomical sites to ‘software’ programmes
- the mechanisms of enduring changes (‘memory’, ‘learning’ and ‘plasticity’) in the CNS are demonstrated
certain activity in the brain – that mental activity is neural activity. The original conjecture by the Greeks – Hippocrates pre-eminent among them – that the brain is the seat of the soul has, it seems, been triumphantly vindicated by modern science: the multiple functions of the secular mind are located in the cerebral cortex. Now, all that remains is to work within this secure framework to tease out the details of what happens in different locations and how those locations relate to one another. And this, we are led to believe, is what has been happening, at an ever-increasing pace, over the last 100 years.

This is very disturbing. It would seem to suggest that our inmost selves are intimately connected with brain activity and, therefore, we humans – far from being metaphysically or ontologically unique beings – are part of the material world. There are, however, many reasons for not jumping straight from the observation that the state of the mind depends on the state of the brain, to the belief that the brain is the seat of the mind.

First, the relationship between objectively observed neural activity and the subjectively experienced contents of consciousness is profoundly puzzling (Fig 1). There are three favoured explanations of the connection between nerve impulses and conscious experiences: the dual aspect theory, the causal interaction theory, and the identity theory. There are obvious flaws with the first two theories, which I have discussed elsewhere. The front runner is the identity theory, which is espoused implicitly or explicitly by most neuroscientists. This theory, which asserts that conscious experiences are identical with certain events located centrally in the brain, actually inherits all the weaknesses of the other two theories but objections to it have focussed on one explanatory gap: that neural activity is not at all like contents of consciousness.

The contrast between the monotonous similarity of neural activity and the infinite variety of the perceived world is one worry. The argument that location in the brain explains all – that sounds are experienced when the hearing neurons are excited and sights are experienced when the visual neurons are excited – is self-evidently circular. Even if there were some way of generating experienced qualities, how would we account for the discrepancy between their variety and the monotonous activity of the nervous system? According to most writers who believe in the identity theory, the basis for the necessary variety is to be found not in the individual impulses but in their patterns, the patterns of large numbers of impulses considered together. They argue that although individual nerve impulses are very much alike, there are millions of different possible patterns of impulses, and it is these patterns that underwrite the infinite variety of the world as presented in subjective experience. The trouble with this argument is that patterns do not exist in, even less for, the elements that make up the pattern. They exist only for an external observer, a consciousness that extracts the pattern.

Consider the array of dots in Fig 2. It could be seen as a single array of nine dots; as an array of six dots on the left and three on the right; or as an array of three dots on the left and six on the right; or as any of a vast number of possibilities. What this tells us is not that the array is infinitely rich in patterns, but that it has no inherent pattern; that its patterns exist only insofar as they are extracted; and they can be extracted only insofar as they are perceived.

What about the more fundamental objection to the identity theory, that nerve impulses not only fail to capture the variety of experience (the differences between different experiences), but they fail to seem like experiences at all? This objection has been countered by an argument from ‘levels’ of description or observation (Fig 3).

Philosophers and some neuroscientists have argued that the relationship between nerve impulses and conscious experiences is like that between water molecules and water. Water molecules are totally unlike water: they do not possess the properties of wetness, shininess, liquidity etc. There is, however, no doubt that water really is identical with H₂O molecules: H₂O molecules and drops of water are the same thing observed at different levels. It is argued, by
analogy, that neural activity and conscious experience are also the same thing, perceived at different levels.

This analogy falls victim to arguments similar to those that undermine the patterns argument. The concept of levels implies levels of observation, and levels of observation presuppose observation and hence consciousness, and so cannot explain the relationship between the seemingly unconscious, third person neural activity of the brain and first person conscious experience. Neither does it explain why some neural activity supposedly has the property of being identical with consciousness while most neural activity – for example that which takes place in the cerebellum, the spinal cord, the peripheral nerves, as well as much of the activity recordable in the cerebral cortex – does not. It would be as though some molecules of H₂O counted as water and others did not.

Nor – and this is absolutely crucial – does it account for the fundamental and unique characteristic of conscious experience, what philosophers call its ‘intentionality’, its character of being about something (Fig 4). One’s consciousness of an object – the mental event, or the nerve impulses, explicitly refers to something other than itself. How do neural discharges in the brain refer back to the object that triggered them? The inward causal chain leading from the object to impulses in the cerebral cortex (represented by the top arrow in the diagram) is consistent with the materialistic framework of neuroscience, but the outward intentional link – whereby the impulses ‘reach out’ to, refer to, are about the seen object – most certainly is not. There is nothing else in nature corresponding to this outward intentional link.

Neural impulses, and hence brain activity and the brain, are even less able to account for the unity of conscious experience. There are impulses all over the brain, but there is no single place where they all come together in a moment of consciousness.

Why do we need to have such a place where it all comes together? This will become evident when we consider the long-range, explicit internal connectedness of consciousness that is necessary for us to be the responsible agents able to operate effectively in our complicated world. Giving this lecture at the College was a commitment that knitted together a multidimensional lacework of moments: the moments many months before it, when I accepted the invitation to speak and discussed the title of my talk; the Sunday mornings in the few weeks preceding it, in which I wrote the lecture; and those moments in which I deployed all sorts of implicit knowledge in order to find my way via taxi and train and foot to the Royal College of Physicians at the right time in the right place, while in the grip of a thousand other preoccupations and floating in a sea of sense data. That I succeeded in arriving to speak as planned is a remarkable tribute to the inexpressibly complex inner organisation of my life and its extendedness across time. Somehow, bursts of electricity in the wetware of the brain seem an inadequate explanation for the exquisitely structured mind that we all have. The problem thrown up by John Locke’s theory of knowledge – that the mind threatens to be a heap of impressions – is not solved by the modularity that phrenologists and contemporary neuroscience attribute to the brain precisely for the reason that while modularity serves the purpose of keeping things tidily apart, it obstructs the need to bring them together in the moment of consciousness.

But there is an even deeper problem than that of bringing everything together. The brain must at the very same time keep vast numbers of projects, actions, micro-projects and micro-actions, distinct. Moreover, to make things even more difficult, those distinct projects must relate to many thousands of others, as each provides the others’ framework of possibility. And worse, moment-to-moment consciousness has to retain a global openness in order that one can enact planned activities in a sea of unplanned contingencies, for example, avoiding a bicyclist while crossing the road.

We tend to overlook the complexity of the most ordinary aspects of our lives when we think about the neurophysiological basis of consciousness. And this is my central message: neuromythology seems halfway plausible only if it is predicated upon a desperately impoverished account of our many-layered, multi-agenda, infinitely complex but wonderfully structured and organised selves. We could summarise the problem very simply as follows. If we try to address the problem of unity of consciousness by adopting a holistic account of the brain, we encounter insuperable difficulties in accounting for the way in which so many different things, which have to be kept apart, are kept apart and do not collapse into mind-mush or delirium. If we try to address the problem of the multiplicity of distinct elements of our conscious lives by adopting a localisationist account of the brain, we encounter equally insuperable difficulties in accounting for the way in which everything comes together sufficiently for us to live active, coherent lives.

The question of unity and control amid diversity – and a continual rain of the half-expected unexpected – picks out a deeper problem: that of accounting for the fact that there is such a thing as the first person (the me, here, now) to
which all this variety is ultimately referred. Without such a unifying element – what Kant called the unity of apperception, and rather unfortunately described as 'the I think that accompanies all my perceptions' – the brain would simply be a colloidal suspension of unhaunted modules, which is how the cognitive scientist seems to present it.

The notion of the first person – the 'I-ness' of consciousness – not only highlights the unity of consciousness necessary for one to act as a responsible agent in a complex world. It opens onto a deeper issue: the origin of the sense of me, here, now; of the suffering agent, the responsible creature who is a viewpoint on the world. And yet this sense that 'I am this thing' is required if my body is to enjoy ownership and I am to have the feeling that I am here now, that I am, to use Heidegger's phrase, a being whose being is an issue for itself. The fact that things matter to one's brain has no basis within the neuroscientific account of the brain/mind. Mattering has no place in the materialist world picture of the identity theorist. In short, there is no basis in the brain either for the unity of consciousness nor for the connection between this unified consciousness and the fundamental intuition of self.

The language of neuromythology

And so we are forced to a conclusion opposite to the one drawn earlier: that consciousness cannot be due to activity in the brain and that cerebral activity is an inadequate explanation of mental activity. Hippocrates' claim that 'from the brain and from the brain only' arise all our experiences cannot be true. And yet this is a message that has certainly not got through to many neuroscientists and their neurophilosophical fellow travellers. How can this be? The reason is this: they increasingly speak to each other and think to themselves in a language that conceals from them the barrenness of their explanations, what I have called the language of neuromythology. This re-describes the mind in mechanical terms and so enables it to remain within the materialist and even biological framework of neuroscience – the mind is construed as a collection of mechanisms. At the same time, it treats these mechanisms as though they were machines which, since the machines are artefacts, may then be spoken of as if they had purposes. Nowadays, the machines in question are usually computers. The mindly brain and the brainy mind merge in the concept of 'information processing', which both are supposed to indulge in.

By this means, the language of neuromythology sidelines the distinctive mystery of human consciousness and forecloses on the ways in which we may think about ourselves and our possibilities.

Separating neuroscience from neuromythology

Hippocrates, instead of asserting that 'men ought to know that from the brain, and from the brain only, arise our pleasures, joys, laughter and jests, as well as our sorrows, pains, grief and tears', ought to have asserted more modestly that a normally functioning brain is a necessary but not a sufficient condition of our experiencing anything at all. This more modest claim would have set us on a different track and not licensed the wild overstatements that neuroscientists and some of their neurophilosophical fellow travellers have made over the last few decades. This is not to say that this escape route is not fraught with difficulties, as I have discussed elsewhere.

One exciting consequence of the failure of neuromythology is that it suggests that making complete sense of the relationship between the brain and consciousness will require a new theory of knowledge or indeed a new account of what kinds of things there are in the world: in short either an epistemological or an ontological rethink. As it is, the neural theory of consciousness, neuromythology, is not only inadequate in itself but depends upon a savagely impoverished account of our own nature as wholly mysterious human animals, at once part of nature and at the same time distant from it, if only to the extent of being able to articulate it.

Meanwhile, neuroscientists may be reassured that what they are doing is not worthless. While they are most certainly not discovering how the mind works, or how it is created in the brain, they are learning more about the conditions under which normal experience and volition is possible: the necessary but not the sufficient conditions. And for me, as a clinician concerned to nurture or encourage those necessary conditions in patients from whom they have been withdrawn, that is good enough. As for neuroscience, metaphysics it is not; worthwhile it certainly is. It just needs to know its limits, so that good science is not discredited by bad philosophy, and scientism does not cause scientists to be justly accused of 'single vision and Newton's sleep'.

References

1 Dennett DC. Consciousness explained. Harmondsworth: Penguin, 1991.
2 Pinker S. How the mind works. Harmondsworth: Penguin 1997.
3 Hippocrates. On the sacred disease. Quoted in: Spillane JD. The doctrine of the nerves. Oxford: Oxford University Press, 1981.
4 Bruyn GW. The seat of the soul. In: Rose FC, Bruyn WF (eds). Historical aspects of the neurosciences. A festskrift for Macdonald Critchley. New York: Raven Press, 1982.
5 Harrington AS. Medicine, mind and the double brain. Princeton University Press, 1987.
6 Peacock A. The relationship between the soul and the brain. In: Rose FC, Bruyn WF (eds). Historical aspects of the neurosciences. A festskrift for Macdonald Critchley. New York: Raven Press, 1982.
7 Tallis R. The explicit animal: a defence of human consciousness (2nd ed). London: Macmillan, 1999.
8 Tallis R. On the edge of certainty. London: Macmillan, 1999.
9 Blake W. Letter to Thomas Butts, 22 November 1802. In: Keynes G. Complete writings. Oxford: Oxford University Press, 1969.

Address for correspondence: Professor R C Tallis, Division of Geriatric Medicine, Clinical Sciences Building, Hope Hospital, Eccles Old Road, Manchester M5 8HD.

E-mail: tallis@fs1.oho.man.ac.uk

Journal of the Royal College of Physicians of London Vol. 34 No. 6 November/December 2000

567