

## A Brief History of the Philosophical Problem of Consciousness

### I. Forms of Consciousness

The term 'consciousness' possesses a huge and diverse set of meanings. It is not even obvious that there is any one 'thing' that all uses of the term have in common which could stand as its core referent (see Wilkes 1988). When we think about consciousness we may have in mind highly complex mental *activities* such as reflective self-consciousness or introspective consciousness, of which perhaps only human beings are capable. Or we may be thinking about something more purely *phenomenal*, perhaps something as apparently simple and unitary as a momentary stab of pain. Paradigm examples of consciousness are the perceptual states of seeing and hearing, but the nature of the consciousness involved is actually complex and far from clear. Are the conscious elements of perception made up only of 'raw' sensations from which we construct objects of perception in a quasi-intellectual operation? Or is perceptual consciousness always of 'completed' objects with their worldly properties?

The realm of consciousness is hardly exhausted by its reflective, introspective or perceptual forms. There is distinctively *emotional* consciousness, which seems to necessarily involve both bodily feelings and some kind of cognitive assessment of them. Emotional states require a kind of evaluation of a situation. Does consciousness thus include distinctive evaluative states, so that, for example, consciousness of pain would involve both bodily sensations and a conscious 'sense' of aversion? Linked closely with emotional states are familiar, but nonetheless rather peculiar, states of consciousness that are essentially other directed, notably empathy and sympathy. We visibly wince when others are hurt and almost seem to feel pain ourselves as we undergo this unique kind of experience.

Philosophers argue about whether all thinking is accompanied by or perhaps even constituted out of sensory materials (images have been the traditional favorite candidate material), and some champion the idea of a pure thought-consciousness independent of sensory components. In any event, there is no doubt that thought is something that often

happens consciously and is in some way different from perception, sensation or other forms of consciousness.

Yet another sort of conscious experience is closely associated with the idea of conscious thought but not identical to it: epistemological consciousness, or the sense of certainty or doubt we have when consciously entertaining a proposition (such as '2 + 3 = 5' or 'the word 'eat' consists of three letters'). Descartes famously appealed to such states of consciousness in the 'method of doubt' (see his *Meditations* 1641/1985).

Still another significant if subtle form of consciousness has sometimes been given the name 'fringe' consciousness (see Mangan 2001, following James 1890, ch. 9) which refers to the 'background' of awareness which sets the context for experience. An example is our sense of orientation or 'rightness' in a familiar environment (consider the change in your state of consciousness when you recognize someone's face who at first appeared a stranger). Moods present another form of fringe consciousness, with clear links to the more 'overtly' conscious emotional states but also clearly distinct from these.

But I think there is a fundamental commonality to all these different forms of consciousness. Consciousness is distinctive for its subjectivity or its 'first-person' character. There is 'something it is like' to be in a conscious state, and only the conscious subject has direct access to this way of being (see Nagel 1974). There is nothing it is like to be a rock, no subjective aspect to an ashtray. But conscious beings are essentially different in this respect. The huge variety in the forms of consciousness makes the problem very complex, but the core problem of consciousness focuses on the nature of subjectivity.

A further source of complexity arises from the range of possible explanatory targets associated with the study of consciousness. One might, for instance, primarily focus on the structure or contents of consciousness. These would provide a valid answer to one legitimate sense of the question 'what is consciousness'. But then again, one might be more interested in how consciousness comes into being, either in a developing individual,

or in the universe at large. Or one might wonder how consciousness, seemingly so different from the purely objective properties of the material world studied by physics or chemistry, fits in with the overall scientific view of the world. To address all these aspects of the problem of consciousness would require volumes upon volumes. The history presented here focuses on what has become perhaps the central issue in consciousness studies, which is the problem of integrating subjectivity into the scientific view of the world (see phenomenology article for discussion of the content and structure aspect of the problem).

## II. The Nature of the Problem

Despite a huge range of diverse opinion, I think it is fair to say that there is now something of a ‘consensus’ view about the origin of consciousness, which I will call here the *mainstream* view. It is something like the following. The world is a purely physical system created some 13 billion years ago in the prodigious event Fred Hoyle labeled the big bang. Very shortly after the big bang the world was in a primitive, ultra-hot and chaotic state in which normal matter could not exist, but as the system cooled the familiar elements of hydrogen and helium, as well as some traces of a few heavier elements, could form. Then very interesting things started to happen, as stars and galaxies quickly evolved, burned through their hydrogen fuel and went nova, in the process creating and spewing forth most of the elements of the periodic table into the increasingly rich galactic environments.

There was not a trace of life, mind or consciousness throughout any of this. That was to come later. The mainstream view continues with the creation of planetary systems. At first these were poor in heavier elements, but after just a few generations of star creation and destruction there were many Earth like planets scattered through the vast – perhaps infinite – expanse of galaxies, and indeed some 7 or 8 billion years after the big bang the Earth itself formed along with our solar system.

We do not yet understand it very well, but whether in a warm little pond, around a deeply submerged hydrothermal vent, amongst the complex interstices of some claylike matrix, as a pre-packaged gift from another world, or in some other way of which we have no inkling, conditions on the early Earth somehow enabled the special – though entirely in accord with physical law – chemistry necessary for the beginnings of life.

But even with the presence of life or proto-life, consciousness still did not grace the Earth. The long, slow processes of evolution by natural selection took hold and ultimately led at some time, somewhere to the first living beings that could *feel* – could feel pain and pleasure, want and fear, could experience sensations of light, sound or odors. The mainstream view sees this radical development as being conditioned by the evolution of neurological behavior control systems in co-evolutionary development with more capable sensory systems. Consciousness thus *emerged* as a product of increasing biological complexity, from non-conscious precursors composed of non-conscious components.

Here we can raise many of the central questions within the problem of consciousness. Imagine we were alien exo-biologists observing the Earth around the time of the emergence of consciousness. How would we know that certain organisms were, while other organisms were not, conscious? What is it about the conscious organisms that explains why they are conscious? Furthermore, the appearance of conscious beings looks to be a development that sharply distinguishes them from their precursors, but the material processes of evolution are not marked by such radical discontinuities. To be sure, we do find striking differences amongst extant organisms. The unique human use of language is perhaps the best example of such a difference, but of course the apes exhibit a host of related, potentially precursor abilities, as do human beings who lack full language use. Thus we have possible models of at least some aspects of our pre-linguistic ancestors which suggest the evolutionary path that led to language.

But the slightest, most fleeting, spark of feeling is a full fledged instance of consciousness which entirely differentiates its possessor from the realm of the non-conscious. Note here a dissimilarity to other biological features. Some creatures have

wings and others do not, and we would expect that in the evolution from wingless to winged there would be a hazy region where it just would not be clear whether or not a certain creature's appendages would count as wings or not. Similarly, as we consider the evolutionary advance from non-conscious to conscious creatures, there would be a range for which we would be unclear about whether or not creatures within that range were conscious or not. But in this latter case, there is a fact whether or not the creatures in that range are feeling anything, however 'dimly' or 'weakly' whereas we do not think there must be a fact about whether a certain appendage is or is not a wing (a dim or faint feeling is 100% a kind of consciousness but a few feathers on a forelimb is not a kind of wing). It is up to us whether to count a certain sort of appendage as a wing or not – it makes no difference, so to speak, to the organism what we call it. But it is not up to us to *decide* whether or not organism X does or does not enjoy some smidgen of consciousness – it either does or it does not.

Lurking behind these relatively empirical questions is a more basic theoretical, or metaphysical, issue. Given that creatures capable of fairly complex behavior were evolving without consciousness, why is consciousness necessary for the continued evolution of more complex behavior? Just as wings are an excellent solution to the problem of evolving flight, brains (or more generally nervous systems) are wonderful at implementing richly capable sensory systems and co-ordinated behavior control systems. But why should these brains be conscious? Although perhaps of doubtful coherence, it's useful to try to imagine our alien biologists as non-conscious beings. Perhaps they are advanced machines well programmed in deduction, induction and abduction. Now, why would they ever posit consciousness in addition to, or as a feature of, complex sensory and behavioral control systems? As Thomas Huxley said: 'how it is that anything so remarkable as a state of consciousness comes about as a result of irritating nervous tissue, is just as unaccountable as the appearance of Djin when Aladdin rubbed his lamp' (1866, 8, 210). We might, rather fancifully, describe this core philosophical question about consciousness as that of how the genie of consciousness gets into the lamp of the brain, or why, to use Thomas Nagel's (1974) famous phrase, there is 'something it is like' to *be* a conscious entity?

### III. Ancient Hints

Of course, the mainstream view has not long been mainstream, for the problem of consciousness cannot strike one at all until a fairly advanced scientific understanding of the world permits development of the materialism presupposed by the mainstream view. A second necessary condition is simply the self-recognition that we are conscious beings possessing a host of mental attributes. And that *has* been around for a long time. Our ancestors initiated a spectacular leap in conceptual technology by devising what is nowadays called ‘folk psychology’. The development of the concepts of behavior explaining states such as belief and desire, motivating states of pleasure and pain, and information laden states of perceptual sensation as well as the complex links amongst these concepts is perhaps the greatest piece of theorizing ever produced by human beings. The power and age of folk psychology is attested by the universal animism of pre-literate peoples and the seemingly ‘innate’ tendencies of very young children to regard various natural or artificial processes as exemplifying agency (see, among many others, Gergeley *et al.* 1995, Perner 1991, Bloom 2004). The persistence of the core mentalistic notions of goal and purpose even in Aristotle’s proto-scientific but highly sophisticated theorizing also reveals the powerful hold these concepts had, and have, on human thought. But to the extent that mentalistic attributes are regarded as ubiquitous, no special problem of relating the mental to the non-mental realm can arise, for there simply is no such realm.

But interesting hints of this problem arise early on in philosophy, as the first glimmerings of a naturalistic world view arise. A fruitful way to present this history is in terms of a fundamental divergence in thought that arose early and has not yet died out in current debate. This is the contrast between emergence and panpsychism. The mainstream view accepts emergence: mind or consciousness appeared out of non-conscious precursors and non-conscious components (note there is both a synchronic and diachronic sense of ‘emergence’). Panpsychism is the alternative view that emergence is impossible and mind must be already and always present, in some sense, throughout the universe (a panpsychist might allow that mind emerges in the trivial sense that the universe may have

been created out of nothing and hence out of ‘non-consciousness’; the characteristically panpsychist position here would be that consciousness must have been created along with whatever other fundamental features of the world were put in place at the beginning). Of course, this divergence transcends the mind-body problem and reflects a fundamental difference in thinking about how the world is structured.

The Presocratic philosophers who flourished some 2500 years ago in the Mediterranean basin were the first in the West to conceive of a something like a scientific approach to nature and it was their conception that eventually led to what we call science. Although their particular theories were understandably crude and often very fanciful, they were able to grasp the idea that the world could be viewed as composed out of elemental features, whose essential characterization might be hidden from human senses and which acted according to constant and universal principles or laws.

The Presocratics immediately recognized the basic dilemma: either mind (or, more generally, whatever the apparently ‘macroscopic’, ‘high-level’, or non-fundamental property at issue) is an elemental feature of the world or it somehow emerges from, or is conditioned by, such features. If one opts for emergence, it is incumbent upon one to at least sketch the means by which new features emerge. If one opts for panpsychism (thus broadly construed) then one must account for the all too obviously apparent total lack of certain features at the fundamental level. For example, Anaxagoras (c. 500-425 BC) flatly denied that emergence was possible and instead advanced the view that ‘everything is in everything’. Anaxagoras explained the obvious contrary appearance by a ‘principle of dominance and latency’ (see Mourelatos 1986) which asserted that some qualities were dominant in their contribution to the behaviour and appearance of things. However, Anaxagoras’s views on mind are complex since he apparently regarded it as uniquely not containing any measure of other things and thus not fully in accord with his mixing principles. Perhaps this can be interpreted as the assertion that mind is ontologically fundamental in a special way; Anaxagoras did seem to believe that that everything has some portion of mind in it while refraining from the assertion that everything has a mind (even this is controversial, see Barnes 1982, p. 405 ff.). On the other hand, Empedocles,

an almost exact contemporary of Anaxagoras, favoured an emergentist account based upon the famous doctrine of the four elements: earth, air, fire and water. All qualities were to be explicated in terms of ratios of these elements. The overall distribution of the elements, which were themselves eternal and unchangeable, was controlled by ‘love and strife’, whose operations are curiously reminiscent of some doctrines of modern thermodynamics, in a grand cyclically dynamic universe. It is true that Empedocles is sometimes regarded as a panpsychist because of the universal role of love and strife (see Edwards 1967 for example) but there seems little of the mental in Empedocles’s conceptions, which are rather more like forces of aggregation and dis-aggregation respectively (see Barnes 1982, pp. 308 ff.).

The purest form of emergentism was propounded by the famed atomist Democritus (c. 460-370 BC). His principle of emergence was based upon the possibility of multi-shaped, invisibly tiny atoms ‘interlocking’ to form an infinity of more complex structures. But Democritus, in a way echoing Anaxagoras and perhaps hinting at the later distinction between primary and secondary properties, had to admit that the qualities of experience (what philosophers nowadays call ‘qualia’, the subjective features of conscious experience) could not be accounted for in this way and chose, ultimately unsatisfactorily, to relegate them to non-existence: ‘sweet exists by convention, bitter by convention, in truth only atoms and the void’. Sorely missed is Democritus’s account of how conventions themselves – the consciously agreed upon means of common reference – emerge from the dancing atoms (thus the ideas of Democritus anticipate the reflexive problem of modern eliminativist materialists (e.g. Churchland 1981) who would enjoin us to consciously accept a view which evidently entails that there is no such thing as conscious acceptance of views – see [Block’s article?](#)).

What is striking about these early struggles about the proper ‘form’ of a scientific understanding of the world, is that the mind and particularly consciousness keep rising as special problems. It is sometimes said that the mind-body problem is not an ancient philosophical issue on the basis that sensations were complacently regarded as bodily phenomena (see Matson 1966), but it does seem that the problem of consciousness was

vexing philosophers 2500 years ago, and in a form redolent of contemporary worries. Also critically important is the way that the problem of consciousness inescapably arises within the context of developing an integrated scientific view of the world.

The reductionist strain in the Presocratics was not favoured by the two giants of Greek philosophy, Plato and Aristotle, despite their own radical disagreements about how the world should be understood. Plato utterly lacked the naturalizing temperament of the Presocratic philosophers, although he was well aware of their efforts. He explicitly criticizes Anaxagoras's efforts to provide naturalistic, causal explanations of human behavior (see *Phaedo*).

Of course, Plato nonetheless has a significant role in the debate since he advances positive arguments in favour of the thesis that mind and body are distinct (see *Phaedo*) as well as providing a basic, and perpetually influential, tri-component based psychological theory (see *Republic*, Book 4). Plato's primary motivation for accepting a dualist account of mind and body presumably stems from the doctrine of the *forms*. These are entities which in some way express the intrinsic essence of things. The form of *circle* is that which our imperfect drawings of circles imitate and point to. The mind can grasp this form, even though we have never perceived a true circle, but only more or less imperfect approximations. The ability of the mind to commune with the radically non-physical forms suggests that mind itself cannot be physical. In the *Phaedo*, Plato (putting words in the mouth of Socrates) ridicules the reductionist account of Anaxagoras which sees human action as caused by local physical events. In its place, the mind is proposed as the final (i.e. teleological) cause of action, merely conditioned or constrained by the physical: 'if it were said that without such bones and sinews and all the rest of them I should not be able to do what I think is right, it would be true. But to say that is because of them that I do what I am doing, and not through choice of what is best – although my actions are controlled by mind – would be a very lax and inaccurate form of expression' (*Phaedo*, 98b ff.).

In general, Plato's arguments for dualism are not very convincing. Here's one. Life must

come from death, because otherwise, since all living things eventually die, everything would eventually be dead. Life can come from death only if there is a distinct 'component', responsible for something being alive, that persists through the life-death-life cycle. That persistent component is soul or mind (*Phaedo* 72c-d). Another argument which Plato frequently invokes (or presupposes in other argumentation) is based on reincarnation. If we grant that reincarnation occurs, it is a reasonable inference that something persists which is what is reincarnated. This is a big 'if' to modern readers of a scientific bent but the doctrine of reincarnation was widespread throughout ancient times and is still taken seriously by large numbers of people. The kernel of a much more powerful argument for dualism lurks here as well, which was deployed by Descartes much later (see below).

Aristotle is famously much more naturalistically inclined than Plato (Raphael's *School of Athens* shows Plato pointing upwards to the heavens while Aristotle gestures downward to Earth as they stare determinedly at each other). But Aristotle's views on mind are complex and obscure; certainly not straightforwardly reductionist (the soul is *not*, for example, a particularly subtle kind of matter such as fire). Aristotle's metaphysics deployed a fundamental distinction between 'matter' and 'form', and any object necessarily instantiates both. A statue of a horse has its matter: bronze, and its form: horse. Aristotle is not using Plato's conception of form here. The form of something is not an other-world separate entity but something more like the *way* in which the matter of something is organized or structured. Nor by 'matter' does Aristotle mean the fundamental physical stuff we refer to by that word; matter is whatever relatively unstructured stuff is 'enformed' to make an object (English retains something of this notion in its use of 'matter' to mean *topic*), so bronze is the matter of a statue, but soldiers would be the matter of an army. Objects can differ in matter but agree in form (two identical pictures, one on paper another on a computer screen) or vice versa. More abstractly, Aristotle regarded *life* as the form of plants and animals and named the form of living things *soul* ('the form of a natural body having life potentially within it' (*De Anima*, bk. 2, ch. 1). Aristotle's views have some affinity both with modern biology's conception of life and the doctrine of psycho-physical functionalism (see Block's

article ?) insofar as he stresses that soul is not a separate thing requiring another ontological realm, but also cannot be reduced to mere matter since the essential attribute is function and organization (for a close and skeptical look at the link between Aristotle's philosophy and modern functionalism see Nelson 1990).

Yet there are elements of Aristotle's account that are not very naturalistic. Early in the *De Anima* Aristotle raises the possibility that the relation between the body and the mind is analogous to that between sailor and ship, which would imply that mind is independent of body. Later Aristotle apparently endorses this possibility where he discusses, notoriously obscurely, the 'active intellect' – the 'part' of the soul capable of rational thought (*De Anima*, bk. 3, chs. 4-5). Aristotle clearly states that the active intellect is separable from body and can exist without it. For Aristotle, like Plato, the problematic feature of mind was its capacity for abstract thought and not consciousness per se, although of course these thinkers were implicitly discussing *conscious* thought and had no conception of mind apart from consciousness.

Discussion of one particular, and highly interesting if perennially controversial, feature of consciousness can perhaps be traced to Aristotle. This is the 'self-intimating' or 'self-representing' nature of all conscious states. Many thinkers have regarded it as axiomatic that one could not be in a conscious state without being aware of that state and Aristotle makes some remarks that suggest he may belong to this school of thought. For example, in Book Three of *De Anima* Aristotle presents, rather swiftly, the following regress argument:

Since we perceive that we see and hear, it is necessarily either by means of the seeing that one perceives that one sees or by another [perception]. But the same [perception] will be both of the seeing and of the colour that underlies it, with the result that either two [perceptions] will be of the same thing, or it [sc. the perception] will be of itself. Further, if the perception of seeing is a different [perception], either this will proceed to infinity or some [perception] will be of itself; so that we ought to posit this in the first instance.

The passage is somewhat difficult to interpret, even in this translation from Victor Caston

(2002) which forms part of an intricate (and controversial) mixed philosophical, exegetical and linguistic argument in favor of the view that Aristotle accepted a self-representational account of conscious states which possessed unique phenomenal properties. Aristotle's argument appears to be that if it is essential to a conscious state that it be consciously apprehended then conscious states must be self-representing on pain of an infinite regress of states each representing (and hence enabling conscious apprehension of) the previous state in the series. The crucial premise that *all* mental states must be conscious is formally necessary for the regress and modern representational accounts of consciousness which accept that conscious states are self-intimating, such as the Higher Order Thought theory, can block the regress by positing non-conscious thoughts which make lower order thoughts conscious by being about them (see Seager 1999, ch. 2, **BLOCK article ?? , THOMPSON / ZAHAVI article ?**).

#### **IV. The Scientific Revolution**

Although the philosophy of the middle ages was vigorous and compendious, the problem of fitting consciousness into the natural world did not figure prominently (for an argument, following in the tradition of Matson 1966, that the medievals' views on the nature of sensation precluded the recognition of at least some versions of the mind-body problem see King 2005). There were many acute studies of human psychology and innovative theoretical work on the content and structure of consciousness and cognition. Of special note is the fourth century philosopher and Church Father, St. Augustine (354-430 BCE). His writings exhibit important insights into the phenomenology of consciousness, especially with regard to the experience of time, will and the self (see especially *Confessions* and *On Free Will*). He was one of the first philosophers to address the problem of other minds, arguing on the basis of introspection and analogy that since others behave as he behaves when he is aware of being in a certain mental state, they too have mental states. In addition, he anticipated certain key features of Descartes's dualistic account of human beings, including Descartes's famous argument from his conscious self-awareness to the certainty of his own existence (*City of God*, Bk. 11, Ch. 21), and the

idea that mind and body, while ontologically entirely distinct, somehow are united in the human person. Here Augustine also broaches one of the key puzzles of Cartesian dualism where he admits the ‘mode of union’ by which bodies and spirits are bound together to become animals is ‘beyond the comprehension of man’ (*City of God*, Bk. 21, Ch. 10). While we also see here that Augustine did not agree with Descartes in denying minds to animals we can also note the complete lack of any idea that this ‘mystery’ poses any special problem for our understanding of the natural world (see O’Daly 1987 for a detailed discussion of Augustine’s philosophy of mind).

In fact, the tenets of Christian dogma, eventually wedded to a fundamentally Aristotelian outlook conspired to suppress any idea that consciousness or mind could be, should be or needed to be explained in naturalistic terms. It was the scientific revolution of the 16<sup>th</sup> and 17<sup>th</sup> centuries that forced the problem into prominence.

Galileo’s distinction between primary and secondary properties, crucial for the development of science insofar as it freed science from a hopelessly premature attempt to explain complex sensible qualities in mechanical terms, explicitly set up an opposition between matter and consciousness: ‘I think that tastes, odors, colors, and so on are no more than mere names so far as the object in which we place them is concerned, and that they reside only in the consciousness. Hence if the living creature were removed all these qualities would be wiped away and annihilated’ (1623/1957, 274). The welcome consequence is that if there are therefore no colors in the world then science is free to ignore them. That was perhaps good tactics in Galileo’s time, but it was a strategic time bomb waiting to go off when science could no longer delay investigating the mind itself.

The mind-body problem in its modern form is essentially the work of a single genius, René Descartes (1596-1650), who reformed the way we think about mind and consciousness, leaving us with a set of intuitions that persist to this day. To take just one topical example, the basic idea behind the fictional technology of the *Matrix* films is thoroughly Cartesian: what we experience is not directly related to the state of the environment, but is instead the result of a complex function – involving essential sensory

and cognitive mediation based upon neural systems – from the environment to our current state of consciousness. Thus two brains that are in identical states ought to be in the same state of consciousness, no matter what differences there are in their respective environments. It now seems intuitively obvious that this is correct (so contemporary philosophers make exotic and subtle arguments against it) and that, to take another stock philosophical example, a brain in a vat, if kept alive in an appropriate chemical bath and if fed proper input signals into its severed nerve endings (cleverly coupled to the output of the brain's motor output nerves) would have experiences which could be indistinguishable from, say, those you are having at this very moment. This thought experiment reveals another of the reformations of philosophy instituted by Descartes: the invention of modern epistemology, for how could you *know* that you are not such a brain in a vat.

Descartes was of course also one of the creators of the scientific revolution, providing seminal efforts in mathematics and physics. But he also saw with remarkable prevision the outlines of neuro-psychology. With no conception of how the nervous system actually works and instead deploying a kind of hydraulic metaphor, Descartes envisioned nerve based sensory and cognitive systems and a kind of network structure in the brain, even – anticipating Hebb – suggesting that connections in the brain are strengthened through associated activation. His notorious discussion of animals as ‘machines’ can be seen as the precursor of a materialist account of cognition.

But Descartes is most remembered and reviled for his insistence upon the strict separation of mind and body which, we are enjoined to believe, required sundering the world itself into radically distinct realms, fundamentally splitting human beings from nature (including their own), denigrated emotion in favour of reason and inspired lack of respect for animals and nature in general. Why was Descartes a dualist? Some have suggested that Descartes lacked the courage to follow his science to its logical and materialist conclusion (the fate of Galileo is said to have had a strong effect on him, or it may be that Descartes really had no wish to harm the Catholic church). But Descartes did have *arguments* for his dualism, some of which still have supporters. They also set out

one of the basic strategies of anti-materialism.

To show that mind and body are distinct it will suffice to show that mind has some property that matter lacks. The general principle here, which is that of the alibi, was codified by another 17<sup>th</sup> century philosopher, Gottfried Leibniz (1646-1716), and is now known as Leibniz's Law: if x has a property which y lacks then x and y are not identical. Descartes argued, for example, that while matter is extended in space, mind takes up no space at all. Thus they could not be identical. It certainly does seem odd to ask how many cubic cm my *mind* takes up (does a broad mind take up more space than a narrow one). But it is not obvious that this is anything more than merely a feature of the conventional way we think about minds. An analogy would be an argument that machines cannot think because they are not *alive*; there is no particular reason to think that the heretofore constant and evident link between life and thought represents anything more than a kind of accident in the way minds happened to be created. In any event, this *strategy* is still at the core of the problem of consciousness. One current line of argument, for example, contends that consciousness has a kind of 'first person subjectivity' (the 'what it is like' to experience something) whereas matter is purely 'third person objective' – hence consciousness and matter must be fundamentally different phenomena.

Descartes, in the sixth of his *Meditations* (1641/1985), also invented an astonishingly novel kind of argument for dualism. The argument is couched in theological terms, but that was merely for purposes of clarity and forcefulness (in the 17<sup>th</sup> century, using God to explain one's argument was impeccable rhetoric). Descartes asked us to consider whether it was at least *possible* that God could destroy one's body while leaving one's mind intact. If it was possible then of course God could perform the feat if He wished. But nothing can be separated from itself! So if it is merely possible that God could sunder mind from body, then they must already be different things. So, anyone who thinks that, say, a persisting consciousness after bodily death is even so much as a bare possibility already thinks that the consciousness is not a physical phenomenon. This argument is valid, but it has a little flaw: how do we know that what we think is possible is truly so? Many are the mathematicians labouring to prove theorems which will turn out to be

unprovable (think of the centuries long effort to square the circle) – what do they think they are doing? Nonetheless, it is a highly interesting revelation that the mere *possibility* of dualism (in the sense considered here) entails that dualism is true.

Cartesian dualism also included the doctrine of mind-body interaction. This seems like common sense: when someone kicks me that causes me to feel pain and anger, and then it is my anger that makes me kick them back. Causation appears to run from body to mind and back again. But as soon as Descartes propounded his theory of mind, this interaction was seen to be deeply problematic. One of Descartes aristocratic female correspondents, the Princess Elisabeth of Palatine, asked the crucial question: ‘How can the soul of man determine the spirits of the body, so as to produce voluntary actions (given that the soul is only a thinking substance)’ (from a letter of May 1643). It’s a fair question and Descartes’s only answer was that the mind-body union was instituted and maintained by God and was humanly incomprehensible. The Princess allowed herself less than *fully* satisfied with this reply.

It was also noticed that Descartes’s dualism conflicted with the emerging understanding of the conservation of certain physical quantities. Descartes himself only accepted that the total amount, but not direction, of motion was conserved. Thus the mind’s ability to wiggle the pineal gland (where Descartes posited the seat of the soul) would redirect motion without violating natural law. But it was soon discovered that it was momentum – or directed motion – that is conserved and thus the mind induced motion of the pineal gland would indeed contradict the laws of nature (one might try to regard this as a feature rather than a bug, since at least it makes Descartes’s theory empirically testable in principle).

In addition to the ontological aspect of his views, Descartes had some interesting insights into the ‘phenomenological’ side of consciousness. For Descartes, the elements of conscious experience are what he called ‘ideas’ (Descartes pioneered the modern use of this term to stand for mental items), and every idea possesses two kinds of ‘reality’: formal and objective. The formal reality of something is simply what it is in itself,

whereas the objective reality is what, if anything, it represents (so, the formal reality of a picture of a horse is paper and paint, a horse is the objective reality). Though Descartes is often pilloried as one who believed that we are only ever conscious of our own ideas, it is far from clear that this is Descartes's position. It is possible to read him instead as a precursor of modern representational theories of consciousness [see Block ??], in which it is asserted that while consciousness essentially involves mental representation, *what* we are conscious of is not the representations themselves but their content (rather in the way that although we must use words to talk about things, we are not thereby always talking *about* words). Descartes says that 'there cannot be any ideas which do not appear to represent some things ...' (*Meditation 3*) and perhaps this suggests that even in cases of illusion Descartes's view was that our experience is of the representational content of the ideas and that we do not, as it were, see our own ideas.

Finally, since Descartes is often misrepresented as denigrating bodily feelings and emotions in favour of 'pure reason' it is worth pointing out that Descartes developed a sophisticated account of the emotions which stresses both their importance and the importance of the bodily feelings which accompany them (1649/1985). Descartes – perhaps contra Aristotle – strenuously denied that the mind was 'in' the body the way a pilot is in a ship, for the intimate connection to the body and the host of functionally significant feelings which the body arouses in the mind in the appropriate circumstances meant that the mind-body link was not a mere communication channel. Descartes declared instead that the mind and body formed a 'substantial union' and that emotional response was essential to cognition.

Despite the fact that *if* one is willing to endorse a dualism of mind and body then Descartes's interactive version seems to be the most intuitively reasonable, the difficulties of understanding how two such entirely distinct realms could causally interact created an avid market for alternative theories of the mind-body relation. Two broad streams of theory can be discerned, which I'll label, not altogether happily, *idealist* and *materialist*. Idealists regard mind or consciousness as the fundamental existent, deny the independent existence of the material world, its apparent reality to be explained as a

function of mentality. Materialists cannot follow such a direct route, for they have great difficulty in outright *denying* the existence of mind and generally content themselves with in some way identifying it with features of matter. The asymmetry in these positions is interesting. Idealists can easily assert that the material world is all illusory. Materialists fall into paradox if they attempt the same strategy – for the assertion that mind is illusory presupposes the existence of illusions, which are themselves mental entities. For a long time (centuries I mean) the idealist position seemed dominant, but the materialists, like the early mammals scrabbling under the mighty dinosaurs, were to have their day.

Early materialists had to face more than an intellectual struggle, since their doctrine stood in clear contradiction with fundamental beliefs endorsed by the Christian churches, and many thinkers have been charged with ‘softening’ their views to avoid ecclesiastical censure. One such is Pierre Gassendi (1592-1655) who espoused an updated version of ancient Epicurean atomism, but who added immortal and immaterial souls to the dance of the atoms. The souls were responsible, in a familiar refrain, for our ‘higher’ *intellectual* abilities. On the materialist core of such a view, nature is ultimately composed of tiny, indivisible and indestructible physical particles whose interactions account for all the complexity and behaviour of organized matter. Gassendi asserted that the ‘sentient soul’, as opposed to the immaterial ‘sapient soul’, was a material component of animals and humans, composed of an especially subtle, quick moving type of matter which is capable of forming the system of images we call imagination and perception (Gassendi also endorsed the empiricist principle that all ideas are based on prior sensory experience). These are literally little images in the brain. Of course, there is a problem here: who is looking at these images? What good does it do to postulate them? For Descartes, the experience of sensory perception or imagination is similarly dependent upon corporeal imagery, but since the visual experience is a mental act, there really is someone to appreciate the brain’s artwork. (Descartes in fact tried to use the imagistic quality of certain experiences as an argument for the existence of material objects, since real images need a material substrate in which they are realized – but Descartes concluded that this argument was far from conclusive.)

A subtle distinction here may have directed philosophers' thinking away from this worry. This is the difference between what are nowadays called substance and property dualism. Descartes is a substance dualist (hence also a property dualist but that is a rather trivial consequence of his view). Substance in general was understood as that which could exist independently (or perhaps requiring only the concurrence of God). Matter was thus a substance, but properties of matter were not themselves substantial, for properties require substance in which to be instantiated. According to Descartes, mind is a second kind of substance, with, naturally, its own set of characteristically mental properties. Thus one basic form of materialism involves merely the denial of mental substance, and the early materialists were keen to make this aspect of their views clear. But denial of substance dualism leaves open the question of the nature of mental properties or attributes (consciousness can be regarded as a feature of the brain, but is no less mysterious for being labeled a property of a physical object).

The problem is clearer in the work of another early materialist, Thomas Hobbes (1588-1679) who, entranced by the new science inaugurated by Galileo, declared that absolutely everything should be explicable in terms of the motions of matter and the efficient causal interaction of material contact. Eventually coming to consider the mind, Hobbes pursues motion into the brain to account for sensory phenomena: 'the cause of sense is the external body ... which presses the organ proper to each sense ... which pressure, by the mediation of the nerves ... continues inwards to the brain ...' (1651/1998, pt. 1, ch. 1). Hobbes goes out of his way to stress that there is nothing immaterial, occult or supernatural here; there is just the various ways that physical events influence our material sense organs: 'neither in us that are pressed are they anything else but divers motions; for motion produceth nothing but motion' (*op.cit.*). But then Hobbes makes a curious remark: speaking of these 'divers motions' in the brain he says 'but their appearance to us is fancy, the same waking that dreaming'. However, he elsewhere states that 'all fancies are motions within us' (1651/1998, pt. 1, ch. 3). Compounding the confusion he also describes our appetites or motivations as motions but says that pleasure and pain are the appearances of these (1651/1998, pt. 1, ch. 6). It would appear that 'appearance' is Hobbes's term for something like phenomenal consciousness and he

seems to be saying that such consciousness is *caused* by motions in the brain but is not identical to them, which of course flatly contradicts his claim that motion can only produce motion. Though obviously Hobbes is not clear about this problem we might anachronistically characterize him as a ‘substance materialist’ who is also a ‘property dualist’.

In any case, materialism was very far from the generally favoured opinion, and the perceived difficulties of Descartes’s substance dualism led instead to a series of inventive alternatives to interactive substance dualism, the two most important being those of Baruch de Spinoza (1632-1677) and Leibniz. In an austere beautiful if forbidding work, the *Ethics* (1677/1985) Spinoza laid out a theory which perhaps, logically, ought to have been Descartes’s. Spinoza notes that substance is that which exists independently of all other things and thus there can be only one ‘maximal’ substance: God. If that is so then matter and mind can only be features of the God-substance (Spinoza called them attributes and asserted there were an infinite number of them, although we are only aware of two). Spinoza’s theory is an early form of what came to be called ‘neutral monism’, which asserts that mind and matter are mere aspects of some underlying kind of thing of which we have no clear apprehension. Particular material or mental individuals (as we would say) are mere modifications of their parent attributes (so your mind is a kind of short lived ripple in the attribute of mind, your body a small disturbance in the material attribute). The attributes are a perfect reflection of their underlying substance but only in respect of one aspect (very roughly like having both a climatographic and topographic map of the same territory). Thus Spinoza believed that the patterns within any attribute would be mirrored in all the others; in particular, mind and body would be automatically and necessarily synchronized. This explains the apparent linkage between mind and body – both are merely aspects of the same underlying substance – while at the same time preserving the causal completeness of each realm. In the illustrative scholium to proposition seven of book two of the *Ethics* (1677/1985) Spinoza writes: ‘a circle existing in nature and the idea of the existing circle, which is also in God, are one and the same thing ... therefore, whether we conceive nature under the attribute of Extension, or under the attribute of Thought ... we shall find one and the same order, or one and the

same connection of causes ...'. On the downside, Spinoza does have to assume that every physical event has a corresponding mental event and he is thus a kind of panpsychist. Even worse (from a 17<sup>th</sup> century point of view) Spinoza's view is heretical, since it sees God as being literally in everything and thus as a material thing not separate from the world.

Leibniz never wrote down his metaphysical system in extensive detail (he was doubtless too busy with a multitude of other projects, such as inventing calculus, rediscovering binary arithmetic, building the first calculating machines, writing endless correspondence and commentary, not to mention his day job of legal counsel and historian to the Hanoverian house of Brunswick), but his views can be reconstructed from the vast philosophical writings he left us. They can be caricatured, in part, as Spinoza's with an infinite number of substances replacing the unique God substance. These substances Leibniz called monads (see Leibniz 1714/1989). Since they are true substances, and hence can exist independently of any other thing, and since they are absolutely simple, they cannot interact with each other in any way (nonetheless they are created by God, who is one of them – here Spinoza seems rather more consistent than Leibniz). Yet each monad carries within it complete information about the entire universe. What we call space and time are in reality sets of relations amongst these monads (or, better, the information which they contain) which are in themselves radically non-spatial and perhaps even non-temporal (Leibniz's vision of space and time emerging from some more elementary systems of relations has always been tempting, if hard to fathom, and now fuels some of the most advanced physics on the planet).

However, Leibniz does not see the monadic substances as having both mental and material aspects. Leibniz's monads are fundamentally to be conceived mentalistically - they are in a way mentalistic automatons moving from one perceptual or apperceptual state to another, all exactly according to a God imposed pre-defined rule. The physical world is a kind of logical construction out of these mental states; one which meets various divinely instituted constraints upon the relation between those aspects matching what we call 'material objects' with those we call 'states of consciousness' – Leibniz

called this the pre-established harmony and it is his explanation for the appearance of mind-body interaction. So Leibniz's view is one that favours the mental realm, that is, it is at bottom a kind of idealism as opposed to Spinoza's neutral monism. As we shall see Leibniz's vision here had a much greater impact on subsequent philosophy than Spinoza's. An important difference between the two theories is that, unlike Spinoza, Leibniz can maintain a distinction between things that have minds or mental attributes from those that do not, despite his panpsychism. This crucial distinction hinges on the difference between a 'mere aggregate' and what Leibniz sometimes calls an 'organic unity' or an organism. Each monad represents the world – in all its infinite detail – from a unique point of view. Consider a heap of sand. It corresponds to a set of monads, but there is no monad which represents anything like a 'point of view' of the heap. By contrast, your body also corresponds to a set of monads but one of these monads – the so called dominant monad – represents the point of view of the system which is your living body. (There presumably are also sub-unities within you, corresponding to organized and functionally unified physiological, and hence also psychological, sub-systems.) Organisms correspond to a hierarchically ordered set of monads, mere aggregates do not. This means that there is no mental aspect to heaps of sand as such, even though at the most fundamental level mind pervades the universe. One last point: you might wonder why you, a monad that represents every detail of the entire universe, seem so relatively ignorant. The answer depends upon another important aspect of the conception of mentality. Leibniz allows that there are unconscious mental states. In fact, almost all mental states are unconscious and low-level monads never aspire to consciousness (or what Leibniz calls apperception). You are aware, of course, only of your conscious mental states and these represent a literally infinitesimal fraction of the life of your mind, the most of which is composed of consciously imperceptible *petite perceptions* (it is galling to think that somewhere within each of our minds lies the invisible answers to such questions as whether there are advanced civilizations in the Andromeda galaxy, but there it is).

For Leibniz the material world is, fundamentally, a kind of illusion, but one of a very special kind. What Leibniz calls 'well grounded' phenomena are those that are in some

way ‘directly’ represented in every monad. Imagine aerial photographs of downtown Toronto taken from a variety of altitudes and angles. The same buildings appear in each photograph, though their appearance is more or less different. But, for example, sun flares caused by the camera lens will not appear in every picture. The buildings would be termed well grounded, the sun flare an illusion. So Leibniz can provide a viable appearance-reality distinction that holds in the world of matter (though it is tricky, since presumably the illusions of any one monad are actually reflected in all monads – hence the weasel word ‘directly’ above). Nonetheless, it is the domain of consciousness which is fundamental and, in the end, the totality of reality, with the physical world being merely a kind of construction out of the mental.

## **V. The Idealist Turn**

In some way, Leibniz represents the culmination of the tradition of high metaphysics: the idea that reason could reveal the ultimate nature of things, and that this nature is radically different from that suggested by common sense. But his model of the material world as mere appearance was taken to its logical next step by the at least superficially anti-metaphysical Immanuel Kant (1724-1804). In Kant (see especially 1781/1929) we see the beginning of the idealism which in one form or another dominated philosophy for more than a century afterwards.

Once mind is established as the sole reality the problem of consciousness, and all the other traditional problems of relating matter to mind, virtually disappear. The problem that now looks big and important is in a way the inverse of the problem of consciousness: how exactly is the material world which we evidently experience to be constructed out of pure and seemingly evanescent consciousness. Two modes of response to this problem can be traced that roughly divide the thinkers of the British Isles (forgive me for including Ireland here) from those of continental Europe; although the geographic categorization becomes increasingly misleading as we enter the 20<sup>th</sup> century. Very crudely, these modes of idealism can be characterized respectively as phenomenalism (material objects are ‘permanent possibilities of sensation’) and transcendental idealism

(a system of material objects represented in experience is a necessary condition for coherent experience and knowledge).

There were, of course, materialists lurking about in this period, though they were nowhere near the heart of philosophical progress; in fact they were frequently not philosophers at all, and quite a number came from the ranks of intellectually inclined medical doctors. One such was Julien de La Mettrie (1709-1751) who outraged Europe, or at least enough of France to require a retreat to Berlin, with his *L'Homme machine* (1748/1987) (see also the slightly earlier *L'Histoire naturelle de l'âme*). In this brisk polemical work, La Mettrie extends the Cartesian thesis that animals are 'mere' machines to include the human animal. But of note here is the same reluctance to shed all reference to the 'specialness' of the mind that we observed in earlier materialists. La Mettrie is willing to deny that there are immaterial mental substances, but describes matter as having three essential attributes: extension, motion and *consciousness*. In *L'Histoire naturelle de l'âme* (1745), La Mettrie makes the interesting point that the intrinsic nature of matter is utterly mysterious to us, and that the attribution of mental properties to it should be no less strange than the attribution of extension and motion, in the sense that we understand what it is about matter itself that supports extension no better – that is not at all – than we understand how it can or cannot support mental properties. This idea has remained an important, if somewhat marginalized, part of the debate about the relation between mind and matter. Although not always very clear about their own positions, most materialists or quasi-materialists of the period (such as John Toland (1670-1722), Paul-Henri D'Holbach (1723-1789), Joseph Priestly (1733-1804) and Pierre-Jean-George Cabanis (1757-1808)) agreed on the approach that denies substance dualism while allowing that matter may have properties that go beyond motion and extension, prominent among which additions are the various mental attributes.

The tide of philosophy was however running in favour of idealism. A central reason for this was independent of the problem of consciousness but stemmed from the epistemological crisis brought about by the Cartesian philosophy (itself but a partial reflection of the general cultural upheaval occasioned by the scientific revolution).

Descartes had argued that the true nature of the world was quite unlike that apparently revealed by the senses, but that reality could be discovered by the ‘light of reason’. Unfortunately, while everyone took to heart the skeptical challenge to conventional wisdom, Descartes’s positive arguments convinced hardly anybody. The core problem was the disconnection between experience and the material world enshrined in Descartes’s dualism. But what if the material world was somehow really a feature of the realm of consciousness for which we ‘obviously’ have infallible access? For example, suppose, as did George Berkeley (1685-1753), that material objects are nothing but ordered sequences of perceptions. We know by introspection that we have perceptions and that they obey certain apparent laws of succession. Under Berkeley’s identification we thereby *know* that there are material objects and the epistemological crisis is resolved.

On the other side of the English Channel, Kant was investigating the deeper problem of how we could know that our perceptions really do follow law governed patterns which guarantee that they can be ‘interpreted’ in terms of a scientifically explicable material world. Kant accepted Leibniz’s view that all we could possibly have knowledge of are constructions out of subjective experience, and that any distinction between reality and appearance within the realm of perception and scientific investigation would have to be based upon some set of relations holding amongst our experiences. He added the remarkable idea that these relations were a reflection of the structure of the mind itself – concepts of space, time and causation are necessary conditions for the existence of experience of an ‘external world’ and are ‘discovered’ in that world because they pre-exist in the mind. There is no reason at all to suppose that they reflect some deeper reality beyond appearances. But they are a necessary condition for having coherent experience at all and hence will and must be discovered in the world which is a construct out of such experience. Kant called this style of reasoning transcendental argumentation. In one sense, however, Kant was an impure idealist. He allowed for the existence of the ‘thing-in-itself’: the necessarily unknowable, mind independent basis of experience. In this respect, Kant’s philosophy is somewhat like Spinoza’s save of course that Spinoza was fully confident that reason could reveal something of the thing-in-itself. Idealists that followed Kant, such as A. Schopenhauer (1788-1869) and the ‘absolute’ idealist G. Hegel

(1770-1831) and many other continental philosophers, and the later followers of Hegel such as the British philosopher F. Bradley (1846-1924), espoused purer forms of idealism (see Schopenhauer 1819/1966, Hegel 1812;16/1969, Bradley 1897/1996). Kant's hypothesis that it was a 'transcendental' condition for the very possibility of introspectible experience that it be lawfully ordered, led to a huge philosophical industry focused, to put it crudely, on the mind's contribution to the structures we find in the 'external world' (an industry that eventually leads into the postmodern deconstructionist ghetto). But this industry, by its nature, did not face the problem of consciousness as defined here and so is not a main player in the drama of this article.

## **VI. Evolution and Emergence**

Instead, as we now enter the heart of the 19<sup>th</sup> century, two crucial non-philosophical developments transformed the problem. These are the rise of Darwinism in biology and, drawn from the brow of philosophy itself, the beginning of scientific psychology. Above all else, the evolutionary theory of Charles Darwin (1809-1882) promised to unify the simple with the complex, to suggest some way that mere atoms could, guided only by the laws of physics, congregate into such complex forms as plants, animals and even human beings. This led immediately to two deep problems: what is life? and how does matter organized via natural selection acquire consciousness? These are both questions about *emergence*, for it certainly appears, if evolution be true, that life springs forth from the lifeless and consciousness appears in beings evolved from non-conscious ancestors composed of utterly non-conscious parts. The first question led to the vitalism controversy, which bears some analogy to the problem of consciousness. Vitalists contended that there was something more to life than mere material organization: a 'vital spark', or *elan vital*. This view of life and its conflict with any materialist account can be traced back at least to the 17<sup>th</sup> century. Another of our philosophically inclined physicians, Nehemiah Grew (1641-1712), who helped found scientific botany and was secretary of the Royal Society in 1677, quaintly put the problem thus, perhaps not atypically confusing the issues of life and consciousness:

The Variety of the Mixture, will not suffice to produce Life ... Nor will its being

mechanically Artificial. Unless the Parts of a Watch, set, as they ought to be, together; may be said to be more Vital, than when they lye in a confused Heap. Nor its being Natural. There being no difference, between the Organs of Art and Nature; saving, that those of Nature are most of all Artificial. So that an Ear, can no more hear, by being an Organ; than an Artificial Ear would do ... And although we add the Auditory nerves to the Ear, the Brain to the Nerves, and the Spirits to the Brain; yet is it still but adding Body to Body, Art to Subtility, and Engine or Art to Art: Which, howsoever Curious, and Many; can never bring Life out of themselves, nor make one another to be Vital (Grew 1701, 33).

Vitalism flourished in the 19<sup>th</sup> century and persisted into the 20<sup>th</sup>, notably in the writings of Hans Driesch (1867-1941) who had discovered that fragments of sea urchin embryos would develop into normal sea urchins, contrary to then current mechanist theory (indeed it is hard to understand how a few of the *parts* of a machine would go on to operate exactly as did the original whole machine). Vitalists thus assumed there must be some special added feature to living things which accounted for the ability to organize and re-organize even in the face of such assaults. It was the unfortunately delayed development of Mendelian ‘information based’ genetics which suggested the answer to Driesch’s paradox, and led to the successful integration of evolution and heredity.

For our purposes, the decline of vitalism provides a cautionary tale but also highlights an important disanalogy between the problems of life and consciousness. Life was seen to be problematic from the materialist point of view because of what it could do, as in Driesch’s sea urchins. It seemed hard to explain the behavioural capacities of living things in terms of non-organic science. Perhaps conscious beings, as living things, present the same problem. But this difficulty was ultimately swept away with the rise of genetics as an adjunct to evolutionary theory. However, in addition to and independent of the puzzle of behaviour, consciousness has an internal or subjective aspect, which life, as such, utterly lacks. What is especially problematic about consciousness is the question of why or how purely material systems could become such that there is ‘something it is like’ to be them.

Another aspect of Darwinism played directly into the mind-matter debate. Darwin himself, and for a long time all Darwinists, was a committed gradualist and assumed that evolution worked by long and slow accumulation of tiny changes, with no infusions of radically new properties at any point in evolutionary history. Applying gradualism to the mind, Darwin went out of his way to emphasize the continuity in the mental attributes of animals and humans (see Darwin 1874).

Gradualism has its difficulties, which have long been noted and persist to this day in talk of punctuated equilibrium (see Eldredge and Gould 1972) and so-called irreducible complexity (Behe 1998 for example). The evolution of the eye was seen as very hard for evolution to explain even by Darwin himself: ‘To suppose that the eye ... could have been formed by natural selection, seems, I freely confess, absurd in the highest degree’ (1859/1967, 167). Of course, the idea that a fully formed eye could appear as the result of one supremely lucky mutational accident is truly absurd and is not what is at issue here. But Darwin went on to give some basis for how the evolution of the eye was possible, and there are nowadays sophisticated accounts of how complex multi-part organs can evolve, and compelling theories of the evolution of particular organs such as the eye (see e.g. Dawkins 1995).

But as noted above, in the most basic sense of the term, consciousness seems to be an all or nothing affair. No non-conscious precursor state seems to give the slightest hint that consciousness would be its evolutionary successor. The tiniest spark of feeling, the weakest and most obscure sensation, are fully states of consciousness. Thus the emergence of consciousness at some point in evolutionary history appears to be an intrusion of true novelty at odds with the smoothly evolving complexity of organisms. William Clifford (1845-1879), a tragically short lived philosophical and mathematical genius (he anticipated general relativity’s unification of gravity with geometry and predicted gravitational waves), put the problem thus:

‘... we cannot suppose that so enormous a jump from one creature to another should have occurred at any point in the process of evolution as the introduction of a fact entirely different and absolutely separate from the physical fact. It is impossible for

anybody to point out the particular place in the line of descent where that event can be supposed to have taken place' (1874/1886, 266).

So, although Darwinism provided great support for and impetus to the materialist vision of the world, within it lurked the old, and still unresolved, problem of emergence.

Perhaps it was time to tackle the mind directly with the tools of science. During the 19<sup>th</sup> century psychology broke away from philosophy to become a scientific discipline in its own right. Despite the metaphysical precariousness of the situation, no one had any doubt that there was a correspondence between certain physical states and mental states, and that it ought to be possible to scientifically investigate that correspondence. The pseudo-science of phrenology was founded on reasonably acceptable principles by Franz Gall (1758-1828) with the aim of correlating physical attributes of the brain with mental faculties, of which Gall, following a somewhat idiosyncratic system of categorization, counted some two dozen, including friendship, amateness and acqistiveness. True, the categorization used is quaint and bizarrely 'high-level' and Gall's shortcut methodology of inferring brain structure from bumps on the skull dubious (to say the least), but the core idea retains vigorous life in today's brain imaging studies and the theory of mental/brain modules. As D. B. Klein said, 'Gall gave wrong answers to good questions' (Klein, p. 669).

Throughout the 19<sup>th</sup> century, one of the primary activities of psychological science and even the main impetus for its creation lay in discovering correlations between psychological states and physical conditions, either of the environment of the subject or brain anatomy discovered via postmortem investigation (but not of course brain *states* which were entirely inaccessible to 19<sup>th</sup> century science). Unlike in the quackery of phrenology, genuine and profound advances were made. Following on foundational work on the physical basis of sensation by Hermann Helmholtz (1821-1894), also famed for introducing the hypothesis that unconscious inference accounts for many aspects of cognition and perception, important discoveries included the connection between certain brain regions and linguistic ability in the work of Paul Broca (1824-1880), seminal studies of stimulus strength and introspected intensity of sensation by Gustav Fechner

(1803-1887) who coined the phrase ‘psycho-physical law’, and the creation of the first psychological laboratory devoted to such studies by Wilhelm Wundt (1832-1920) who also developed the first distinctive research methodology of psychology – that of introspectionism.

From the point of view of the problem of consciousness these developments point to a bifurcation in the issue. Almost all the thinkers associated with the birth of psychology endorsed some form of idealism as the correct metaphysical account of mind and matter; and none of prominence were materialists. They were nonetheless keen on studying what we would call the neural bases of consciousness, and never questioned the legitimacy of such studies. It is useful to distinguish the study of the ‘structure’ of consciousness from the question of the ultimate nature of consciousness and its place in the natural world. The pioneers of psychology were, so to speak, officially interested in the structure of consciousness, both its introspectible experiential structure and its structural link to physical conditions (both internal and external to the body).

The growth of interest in these questions can also be seen in more purely philosophical work in the rise of the phenomenological movement, although of course the philosophers were not particularly interested in investigating correlations between mental and material conditions but rather focused on the internal structure of ‘pure consciousness’.

Phenomenology was foreshadowed by Franz Brentano (1838-1917) who in a highly influential work, *Psychology from the Empirical Standpoint* (1874/1973, pp. 121 ff.), offered an influential statement of the view that mental states were self-intimating coupled with an updated version of Aristotle’s regress argument (Brentano rather generously credits Aristotle for his whole line of thought here).

Brentano also reminded philosophers of a feature of mental states which had almost been forgotten since it had first been noted in the middle ages (though Descartes’s notion of ‘objective reality’ is closely related). Brentano labeled this feature *intentionality*, which is the ‘directedness’ or ‘aboutness’ of at least many mental states onto a content, which may or may not refer to an existing object. If I ask you to imagine a unicorn, you are easily

able to do so, despite the fact that there are no unicorns. Now, what is your thought about? Evidently not any real unicorn, but neither is your thought about the image of a unicorn produced in your imagination or even just the idea of a unicorn. For if I asked you to think about your image or idea of a unicorn you could do that as well, but it would be a different thought, and a rather more complex one. One way to think about this is to say that any act of imagination has a certain representational content, and imagining a unicorn is simply the having of a particular unicorn-content (in the 'appropriate' way as well, for imagination must be distinguished from other content bearing mental acts). The consciousness involved in such an act of imagination is the presentation of that content to your mind. This is not to say that you are aware of your mental state whenever you imagine but rather it is through having such a state that you are conscious of what the state represents, although Brentano himself held that any conscious state presented *itself* as well as its content to the subject. The failure to notice the intentional aspect of consciousness had bedeviled philosophy, leading to a plethora of theories of thought and perception that left us in the awkward position of never being aware of anything but our own mental states.

Brentano went so far as to declare intentionality *the* mark of the mental, the unique property that distinguished the mental from the physical. Of course, many other things, such as pictures, words, images on television, electronic computation, etc. have representational content, but arguably these all get their content derivatively, via a mental interpretation. Uniquely mental or not, intentionality poses an extremely difficult question: how is it that mental states (or anything else) can acquire representational content? Perhaps if one accepts, as so many of the thinkers of this period did, that mind is the bedrock reality, then one can accept that it is simply a brute fact, an essential property of mentality, that it carry representational content. No explanation of this basic fact can be given in terms of anything simpler or more fundamental.

However, if one aspires to a materialist account of mind then one cannot avoid this issue. A frequent metaphor which materialists of the time appealed to was that of biological secretion, perhaps first explicitly articulated by Cabanis in his *Rapports du physique et*

*du moral de l'homme* (1802) who proclaimed that the brain secretes thought as the liver secretes bile. As it stands, this is little more than a declaration of loyalty to the materialist viewpoint, for we expect there should be an explication of the process and nature of such secretion just as there is such an account of the production of bile. Just which part of the brain generates these 'secretions' and how do they manage to possess representational or phenomenal content? Nonetheless, the metaphor was effective. It was approved by Darwin himself (a closet materialist), who (privately) endorsed its repetition by John Elliotson (1791-1868) – physician, phrenologist, mesmerist and the so-called strongest materialist of the day (see Desmond and Moore 1994, 250 ff.). In one of his private notebooks Darwin modified the metaphor in an interesting way, writing 'Why is thought, being a secretion of brain, more wonderful than gravity as a property of matter'? This is striking because it clarifies how the metaphor implicitly suggests that it is a brute fact that brains produce thought, just as it is a brute fact that matter is associated with gravitation. Note also how the power to gravitate seems remote from matter's core properties of extension, exclusion and mass, and, at least in the Newtonian view, provides the almost miraculous ability to affect all things instantaneously at a distance. Nonetheless the essential emptiness of the metaphor did not go unremarked. William James (1842-1910) wrote 'the lame analogy need hardly be pointed out ... we know of nothing connected with liver and kidney activity which can be in the remotest degree compared with the stream of thought that accompanies the brain's material secretions' (1890/1950, pp. 102-3).

Leaving aside once again this metaphysical issue, workers focused on the structure and meaning of the contents of consciousness along with their empirically determinable relationship to a host of internal and external material conditions. I have referred to early scientific psychology above, but I would also put Sigmund Freud (1856-1939) in this group. Although an advocate of materialism, his theory of mind was focused on psychological structure rather than explications of how matter gives rise to mind. In philosophy, this emphasis eventually led to the birth of a philosophical viewpoint explicitly dedicated to investigating the 'inner structure' of consciousness: the phenomenology of Edmund Husserl (1859-1938). In the newly scientific psychology

under the guidance of Wundt, introspection became the paradigm research methodology, raising fundamental questions such as whether all thought was necessarily accompanied by, or even constituted out of, mental imagery. Unfortunately, this methodology suffered from inherent difficulties of empirical verification and inter-observer objectivity, which eventually brought it into disrepute, probably, overall, to the detriment of scientific psychology.

Though James decried the gross metaphor of consciousness as a brain secretion, he introduced one of the most potent and durable metaphors, that of the *stream of consciousness*. In his remarkably compendious work, *The Principles of Psychology*, which remains to this day full of fresh insight, James devoted a chapter to the stream of thought in which he noted that on-going consciousness is continuous, meaning ‘without breach, crack or division’ (1890/1950, 237) and that, by contrast, ‘the breach from one mind to another is perhaps the greatest breach in nature’. James of course allowed that there were noticeable gaps in one’s stream of consciousness, but these are peculiar gaps such that we sense that both sides of the gap belong ‘together’ in some way, he also noted that the stream is a stream of *consciousness*, and that *unnoticed* temporal gaps – which are perfectly conceivable – are simply not part of the stream. Throughout his writings James’s exhibits a keen and durable interest in the structure and contents of the stream of consciousness, even delving enthusiastically into mystical and religious experience.

Along with virtually all psychological researchers of the time, James was no materialist. His metaphysics of mind is complex and somewhat obscure, wavering between a neutral monism and a form of panpsychism. James heaped scorn (and powerful counter arguments) upon crude forms of ‘molecular’ panpsychism, what he called the ‘mind dust’ theory (see 1890/1950, ch. 5), but his monism leaned decidedly towards the mental pole. In a notebook he wrote that ‘the constitution of reality which I am making for is of the psychic type’ (see Cooper 1990).

This unclarity may arise from the epistemological asymmetry between our apprehension of mind and matter. We seem to have some kind of direct access to the former – when we

feel a pain there is an occurrence at least some properties of which are made evident to us. We do not seem to have any similarly direct awareness of the nature of matter. Thus the avowed neutrality of neutral monism tends to slide towards some kind of panpsychism. From another point of view, the asymmetry encourages the association of some forms of phenomenalism with neutral monism.

For example, the highly influential British philosopher John Stuart Mill (1806-1873) endorsed a phenomenalism which regarded material objects as ‘permanent possibilities of sensation’. This allows for the interposition of a *something we know not which* lurking behind our sensations (what might be called ‘unsensed sensibilia’ – see Mill 1865/1983 and Wilson 2003), but the seemingly unbridgeable gap between this ur-matter and our perceptual experiences creates a constant pressure to replace it with entirely mental sequences of sensations. To be sure, intuition suggests that material objects exist unperceived, but this ‘existence’ can, perhaps, be analysed in terms of dispositions to have certain sensations under certain *mentalistically* defined conditions. Furthermore, since our relation with the unknowable basis of matter is entirely mentalistic why not accept that the primal material is itself mental (a view which can lead either back to idealism or to some form of panpsychism)? Bertrand Russell (1872-1970) devoted great effort to developing Mill’s phenomenalism as a kind of neutral monism (see Russell 1927) in which what we call matter has intrinsic mental properties with which we are directly acquainted in experience – thus Russell’s seemingly bizarre remark that when a scientist examines a subject’s brain he is really observing a part of his *own* brain (for an updated defense of a Russellian position see Lockwood 1991).

His one time collaborator, Alfred North Whitehead (1861-1947) pursued the alternative panpsychist option in a series of works culminating in the dense and obscurely written *Process and Reality* (1929). Roughly speaking Whitehead proposed a radical reform of our conception of the fundamental nature of the world, placing *events* (or items that are more event-like than thing-like) and the ongoing *process* of their creation as the core feature of the world, rather than the traditional triad of matter, space and time. His panpsychism arises from the idea that the elementary events that make up the world

(which he called *occasions*) partake of mentality in some - often extremely attenuated - sense, metaphorically expressed in terms of the mentalistic notions of creativity, spontaneity and perception. Whitehead's position nicely exposes the difficulty in maintaining a 'pure' neutral monism. Matter must have some underlying intrinsic nature. The only intrinsic nature we seem to be acquainted with is consciousness. Thus it is tempting to simplify our metaphysics by assigning the only known intrinsic nature to matter. We thus arrive at panpsychism rather than neutral monism (for an introduction to Whitehead's philosophy of mind see Griffin 1998).

Such high metaphysical speculations, though evidently irresistible, seem far from the common sense view of matter which was more or less enshrined in the 'world view' of 19<sup>th</sup> century science, which began then to fund the rapid and perpetual development of technology we are now so familiar with, and which greatly added to the social prestige of science. If we take the scientific picture seriously – and it came to seem irresponsible not to – then the central mystery of consciousness becomes that of the integration of mind with this scientific viewpoint. This is the modern problem of consciousness, which bypasses both idealist metaphysics and phenomenalist constructionism.

But how could such integration be achieved? An important line of thought begins with some technical distinctions of Mill. In his *System of Logic* (1843/1963) Mill attempted a compendious classification of scientific law, two forms of which he called 'homopathic' and 'heteropathic'. Homopathic laws are ones in which the resultant properties of a system are the mere additive results of the properties of the system's components. For example, the laws of motion are homopathic: the motion of an object is the result of all the forces acting on the object, and the resultant force is simply the vector addition of each separate force. Heteropathic laws are ones in which the resultant properties are *not* simply the sum of the properties of the components. It was George Lewes (1817-1878) – now best remembered as the consort of George Elliot – who coined the term 'emergent' to refer to heteropathic effects (he used 'resultant' to mean those features which Mill called homopathic effects). Here it is important to distinguish the more general notion of homopathic effects from what is sometimes called part-whole reductionism. The latter

may well be false of the world: there are reasonable arguments that some physical properties are non-local and perhaps in some way ‘holistic’ (both general relativity and quantum mechanics can be invoked to support these contentions). But the crucial question about homopathic versus heteropathic effects is whether the fundamental physical state of the world along with the basic physical laws determines all other, higher-level, properties and laws. If not, we have true emergence.

The emergentists postulated that consciousness was a heteropathic effect, or emergent property, of certain complex material systems (e.g. brains). Emergentism may seem no more than extravagant metaphysical speculation, except that at the time it was widely conceded that there were excellent candidate emergent properties in areas other than consciousness. That is, it seemed there were independent grounds for endorsing the concept of emergence, which – thus legitimated – could then be fruitfully applied to the mind-body problem. The primary example of supposedly uncontentious emergent properties were those of chemistry. It was thought that, for example, the properties of water could not be accounted for in terms of the properties of oxygen and hydrogen *and* the laws of nature which governed atomic level phenomena. Emergentists, of which two prominent were Conwy Lloyd Morgan (1852-1936) and C. D. Broad (1887-1971), recognized that the complexity of the interactions of the components of a system could present the appearance of emergence when there was none. Broad liked to imagine a ‘mathematical archangel’ who knew the laws of nature as they applied at the submergent level, knew the configuration of the components, and suffered from no cognitive limitations about deducing the consequences of this information. If the archangel could figure out that water would dissolve salt by considering only the properties of oxygen, hydrogen, sodium and chlorine as well as the laws which governed their interaction at the atomic level, then this aspect of H<sub>2</sub>O would fail to be an emergent property.

Thus the emergentists would have scoffed at current popular examples of emergence such as John Conway’s Game of Life and chaotic dynamical systems. Such examples represent nothing more than ‘epistemological emergence’. Cognitive and physical limitations – albeit quite fundamental ones – on computational power and data

acquisition prevent us (or our machines) from deducing the high-level properties of complex systems, but this is not a metaphysical barrier. The mathematical archangel could figure out the effect of the butterfly's flight on future weather.

But the emergentists believed that the world really did contain non-epistemological emergence; in fact, it was virtually everywhere. They regarded the world as an hierarchical cascade of emergent features built upon other, lower level, emergent features. Unfortunately for them, their linchpin example, chemistry, was the masterpiece of the new quantum mechanics of the 1920s, which basically provided new laws of nature which opened the door – in principle – to the deduction of chemical properties from atomic states (nowadays we even have *de novo* calculation of some simple chemical features based on the quantum mechanical description of atomic components). Of course, this does not *demonstrate* that there is no real emergence in the world, but without any uncontroversial example of it, and with the growing ability of physics to provide seemingly complete accounts of the basic structure of the world, the emergentist position was devastated (see McLaughlin 1992).

On the mainstream view articulated in Section II, emergentism seems no less metaphysically extravagant than the other positions we have considered. Emergentism espouses a form of property dualism and postulates that the novel emergent properties of a system would have distinctive causal powers, going beyond those determined solely by the basic physical features of the system (seemingly courting violation of a number of basic conservation laws).

Not that the science of psychology provided a more palatable alternative. Early in the 20<sup>th</sup> century, the introspectionist methodology as well as the sophisticated sensitivity to the issues raised by consciousness by psychologists such as James disappeared with the rise of a soulless behaviourism that at best ignored the mind and at worst denied its very existence. It took until half way through the 20<sup>th</sup> century before philosophy and psychology grappled with the problem of the mind in new ways. In psychology, the so-called cognitive revolution made appeal to inner mental processes and states legitimate

once again (see Neisser 1967 for a classic introduction to cognitive psychology), although scientific psychology largely steered clear of the issue of consciousness until near the end of the 20<sup>th</sup> century [see article on current psychology/neurology of consciousness?].

In philosophy, the 1950s saw the beginning of a self-conscious effort to understand the mind and, eventually, consciousness as physical through and through in essentially scientific terms. This was part of a broader movement based upon the doctrine of scientific realism, which can be roughly defined as the view that it is science that reveals the ultimate nature of reality rather than philosophy, or any other non-empirical domain of thought. Applied to the philosophy of mind and consciousness, this led to the rise of the *identity theory* (see e.g. Smart 1959, or for a more penetrating and rather prescient account Feigl 1958, [and see Block's article?]), which was a kind of turning point in philosophy, in which it was self-consciously assumed that since physical science should be the basis for our beliefs about the ultimate nature of the world, philosophy's job in this area would henceforth be to show how the mind, along with everything else, would smoothly fit into the scientific picture of the world. It embraced what I earlier called the mainstream view, and began with high optimism that a scientific outlook would resolve not just the problem of consciousness but perhaps *all* philosophical problems. But subsequent work has revealed just how extraordinarily difficult it is to fully explicate the mainstream view, especially with regard to consciousness. In the face of unprecedented expansion in our technical abilities to investigate the workings of the brain and an active and explicit scientific, as well as philosophical, effort to understand how, within the mainstream view, consciousness emerges we find that the ultimate problem remains intractable and infinitely fascinating.

William Seager  
University of Toronto at Scarborough

## Bibliography

- Aristotle (1984). *The Complete Works of Aristotle*. J. Barnes (ed.), Princeton: Princeton University Press.
- Augustine (c. 400/1998). *Confessions*, Oxford: Oxford University Press.
- Augustine (c. 426/1998). *City of God*, Cambridge: Cambridge University Press.
- Barnes, J. (1982). *The Presocratic Philosophers*, London: Routledge and Kegan Paul.
- Behe, M. (1998). *Darwin's Black Box: The Biochemical Challenge to Evolution*, New York: Free Press.
- Berkeley, G. (1710/1998). *A Treatise Concerning the Principles of Human Knowledge*, Oxford: Oxford University Press.
- Bloom, P. (2004). *Descartes's Baby: How the Science of Child Development Explains What Makes Us Human*, New York: Basic Books.
- Bradley, F. (1897/1966). *Appearance and Reality*, 2<sup>nd</sup> Ed. Oxford: Clarendon Press.
- Brentano, C. (1874/1973). *Psychology from an Empirical Standpoint*, London: Routledge and Kegan Paul.
- Broad, C. D. (1925). *The Mind and Its Place in Nature*, London: Routledge and Kegan Paul.
- Cabanis, P. (1802/1981). *Rapports du physique et du moral de l'homme*, Paris: Crapart, Caille et Ravier. In English as *On the Relations between the Physical and Moral Aspects of Man*, G. Mora (ed.) M. Saidi (trans). Baltimore: Johns Hopkins University Press.
- Caston, V. (2002). 'Aristotle on Consciousness', in *Mind*, 111(444), pp. 751-815.
- Clifford, William K. (1874). 'Body and Mind', in *Fortnightly Review*, December. Reprinted in *Lectures and Essays*, Leslie Stephen and Frederick Pollock (eds.), London: Macmillan, 1886.
- Churchland, Paul (1981). 'Eliminative Materialism and Propositional Attitudes', *Journal of Philosophy*, 78, pp. 67-90.
- Cooper, W. E. (1990). 'William James's Theory of Mind', *Journal of the History of Philosophy*, 28, 4, pp. 571-93.
- Darwin, C. (1859/1967). *The Origin of Species by Natural Selection*, London: Dent (Everyman's library).
- Darwin, C. (1874). *The Descent of Man*, London: Crowell.
- Dawkins, R. (1995). *River out of Eden*, New York: Basic Books.

- Descartes, R. (1641/1985). *Meditations on First Philosophy*, in J. Cottingham, R. Stoothoff and D. Murdoch (eds.) *The Philosophical Writings of Descartes*, Cambridge: Cambridge University Press.
- Descartes, R. (1649/1985). *The Passions of the Soul*, in J. Cottingham, R. Stoothoff and D. Murdoch (eds.) *The Philosophical Writings of Descartes*, Cambridge: Cambridge University Press.
- Desmond, A. and J. Moore (1994). *Darwin: The Life of a Tormented Evolutionist*, New York: Norton.
- Edwards, P. (1967). 'Panpsychism', in *The Encyclopedia of Philosophy*, vol. 5, P. Edwards (ed.), New York: Macmillan.
- Eldredge, N. and S. Gould (1972), 'Punctuated Equilibria: An Alternative to Phyletic Gradualism', in T. Schopf (ed.) *Models in Paleobiology*, San Francisco: Freeman Cooper.
- Feigl, H. (1958). 'The Mental and the Physical', in H. Feigl, M. Scriven and G. Maxwell (eds.) *Minnesota Studies in the Philosophy of Science, Vol. 2: Concepts, Theories and the Mind-Body Problem*, Minneapolis: University of Minnesota Press.
- Fechner, G. (1946). *The Religion of a Scientist*, (selections of Fechner's writing in English translation), W. Lowrie (ed./trans.), New York: Pantheon.
- Galileo, G. (1623/1957). *The Assayer*. Reprinted in Stillman Drake (ed.), *Discoveries and Opinions of Galileo*, New York: Doubleday.
- Gergely, G, Z. Nádasy, G. Csibra and S. Biró (1995). 'Taking the intentional stance at 12 months of age', in *Cognition*, 56 (2): 165-193.
- Grew, N. (1701). *Cosmologia Sacra: or a Discourse of the Universe as it is the Creature and Kingdom of God*. London: Rogers, Smith and Walford.
- Griffin, D. (1998). *Unsnarling the World Knot: Consciousness, Freedom and the Mind-Body Problem*, Berkeley: University of California Press.
- Hegel, G. (1812;16/1969). *The Science of Logic*, A. Miller (trans.), London: Allen and Unwin.
- Hobbes, T. (1651/1998). *Leviathan*, J. Gaskin (ed.), Oxford: Oxford University Press.
- Holbach, Baron d' (1770/1970). *Système de la nature; ou, Des lois du monde physique et du monde moral* (published under the pseudonym J. Mirabaud). In English as *The System of Nature: or, Laws of the Moral and Physical World*, H. Robinson (trans.), New York: Burt Franklin.
- Husserl, E. (1931/1960). *Cartesian Meditations*, D. Cairns (trans.), The Hague: Martinus Nijhoff.
- Huxley, T. (1866). *Lessons in Elementary Physiology*, London: Macmillan.
- James, W. (1890/1950). *The Principles of Psychology*, v. 1, New York: Henry Holt and

Co. Reprinted in 1950, New York: Dover.

Kant, I. (1781/1929). *Critique of Pure Reason*, N. Kemp Smith (ed./trans.), New York: St. Macmillan.

King, P. (2005). 'Why Isn't the Mind-Body Problem Mediaeval?' in H. Lagerlund and O. Pluta (eds.) *Forming the Mind – Conceptions of Body and Soul in Late Medieval and Early Modern Philosophy*, Berlin and New York: Springer Verlag.

Klein, D. (1970). *A History of Scientific Psychology*, New York: Basic Books.

La Mettrie, J. (1745). *L'histoire naturelle de l'âme*, The Hague.

La Mettrie, J. (1748/1987). *L'homme machine / Man a Machine*, La Salle: IL, Open Court.

Leibniz, G. (1714/1989). *Monadology*, in R. Ariew and D. Garber (eds./trans.) *G. W. Leibniz: Philosophical Essays*, Indianapolis: Hackett.

Lockwood, M. (1991). *Mind, Brain and the Quantum*, Oxford: Blackwell.

McLaughlin, B. (1992). 'The Rise and Fall of British Emergentism', in A. Beckermann, H. Flohr and J. Kim (eds.), *Emergence or Reduction?*, Berlin: De Gruyter.

Mangan, B. 'Sensation's Ghost: The Non-Sensory "Fringe" of Consciousness', *Psyche* 18, no. 7 (2001).

Matson, W. (1966). 'Why Isn't the Mind-Body Problem Ancient?', in P. Feyerabend and G. Maxwell (eds.) *Mind, Matter and Method: Essays in Philosophy and Science in Honor of Herbert Feigl*, Minneapolis: University of Minnesota Press.

Mill, J. (1843/1963). *A System of Logic*, volumes 7 and 8 of J. Robson (ed.) *Collected Works of John Stuart Mill*, Toronto: University of Toronto Press.

Mill, J. (1865/1963). *An Examination of Sir William Hamilton's Philosophy*, volume 9 of J. Robson (ed.) *Collected Works of John Stuart Mill*, Toronto: University of Toronto Press.

Morgan, C. (1923). *Emergent Evolution*, London: Williams and Norgate.

Morton, Peter (1997). *A Historical Introduction to the Philosophy of Mind*, Peterborough, ON: Broadview Press.

Mourelatos, Alexander (1986). 'Quality, Structure, and Emergence in Later Pre-Socratic Philosophy', *Proceedings of the Boston Area Colloquium in Ancient Philosophy*, 2, 127-194.

Nagel, T. (1974). 'What is it Like to be a Bat?' in *Philosophical Review*, 83, 435-50.

Neisser, U. (1967). *Cognitive Psychology*, New York: Appleton-Century-Crofts.

Nelson, J. (1990). 'Was Aristotle a Functionalist?', *Review of Metaphysics*, 43, 791-802.

- Nussbaum, M. C. & Putnam, H. (1992). 'Changing Aristotle's mind', in M. Nussbaum and A. Rorty (eds.), *Essays on Aristotle's De Anima*, Oxford: Clarendon Press, 27-56.
- O'Daly, G. (1987). *Augustine's Philosophy of Mind*. London: Duckworth.
- Perner, J. (1991). *Understanding the Representational Mind*, Cambridge, MA: MIT Press.
- Plato (1961). *Plato: Collected Dialogues*, E. Hamilton and H. Cairns (eds.), Princeton: Princeton University Press.
- Priestley, J. (1777/1975). *Disquisitions Relating to Matter and Spirit*, London: Printed for J. Johnson. Reprint New York: Arno Press.
- Russell, B. (1927). *The Analysis of Matter*, London: Kegan Paul.
- Schopenhauer, A. (1819/1966). *The World as Will and Representation*, E. Payne (trans.), New York: Dover Books.
- Seager, W. (2001). 'Panpsychism', in the *Stanford Encyclopedia of Philosophy*, E. Zalta (ed.), URL = <http://plato.stanford.edu/archives/sum2001/entries/panpsychism>.
- Smart, J. (1959). 'Sensations and Brain Processes', in *Philosophical Review*, 68, 141-56. Reprinted in slightly revised form in V. Chappell (ed.) *The Philosophy of Mind*, Englewood Cliffs, NJ: Prentice-Hall.
- Spinoza, B. (1677/1985). *Ethics*, in E. Curley (ed./trans.) *The Collected Works of Spinoza, Vol. 1*, Princeton: Princeton University Press.
- Toland, J. (1704). *Letters to Serena*, London: Printed for Bernard Lintot.
- Whitehead, A. (1929). *Process and Reality: an Essay in Cosmology*, New York: Macmillan.
- Wilkes, K. (1988). '\_\_\_\_\_, Yishi, Duh, Um, and Consciousness', In A. Marcel and E. Bisiach (eds.) *Consciousness in Contemporary Science*, Oxford: Oxford University Press.
- Wilson, F. (2003). 'John Stuart Mill', in the *Stanford Encyclopedia of Philosophy*, E. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2003/entries/mill>.
- Wundt, W. (1892/1894). *Vorlesungen über die Menschen- und Thierseele*, Hamburg: L. Voss. An English translation is available as *Lectures on Human and Animal Psychology*, J. Creighton and E. Titchener (trans.), London: S. Sonnenschein.