

Theories of Mind

Within our discussion of consciousness and artificial intelligence, a working knowledge of the main theories of mind would be useful – not least because they will help in terms of making the link between the physical processes of the brain, our actions and thoughts, and how we conceive of consciousness.

Theories of mind have a long history – Aristotle, for instance, considered our faculty of reason (or “intellect”) to be an aspect of the soul; I will not spend too much time on this, but it is worth outlining his conception of soul.

Aristotle believed the soul to be divided into two main elements: the rational part and irrational part. These are then subdivided into four further parts – two parts for each element.

The rational element of the soul contains the *scientific part* (that concerned with knowledge of facts and how we come to know them) and the *calculative part* (the part concerned with deliberation and judgement).

The irrational element of the soul contains the *desiderative part* (what we want and desire e.g. sex, chocolate, alcohol etc) and the *vegetative part* (what we need to be healthy and the instincts we have to obtain it).

These aspects interact and allow us to obtain wisdom. Put differently: we work with the facts and from them calculate how we should act – finding the mean between rashness and cowardice, for example; we also work with the irrational part of our soul and, through the rational part, work with our wants, desires, needs and instincts. Through these interactions we gradually develop our character.

How does this relate to consciousness? Aristotle tends to equate thinking with perception but this has problems. Firstly, does a being have to be able think in order to be conscious? – Can thinking occur without consciousness? Secondly, although Aristotle distinguished between perception and thought insofar as they involve different faculties, they are largely similar in terms of how they are modelled – perception requires the reception of the form of the world by the appropriate faculty – the form of sound (hearing) requires the ear, for example – thinking requires the receiving of an appropriate form by the intellect and so on. Thinking, according to Aristotle, takes place when the mind is shaped by the object that affects it – in other words, it becomes like its object(s) in the sense that it is corresponding or similar in its form and relations to the object(s) it apprehends. But what does this really tell us about consciousness, thought or intelligence? I will leave this for you to think about (perhaps we will have discussed it in the class).

After Aristotle, conceptions of mind and thought became largely answerable to religious ideas about the nature of the soul. The soul was largely conceived as an immaterial entity in which inhered some form of divine presence. This idea was fleshed out and developed by Rene Descartes in his *Meditations on First Philosophy* – this work, you will remember, contains his famous words, “I think therefore I am.” However, it is also where Descartes develops the idea of **Substance Dualism** (more often known as **Cartesian Dualism**).

Cartesian (Substance) Dualism:

Separates the mental from physical i.e. it separates such things as thoughts and dreams from brains and limbs (the former being immaterial, the latter physical).

Descartes believed mind and body to be two separate substances – anything physical that takes up space is an *extended substance* while the mind is unextended in that it is entirely non-spatial and not part of the physical world.

This view is particularly appealing to those who have notions about a transcendental soul. Those who advocate the Cartesian position claim that there is a significant body of evidence to support the view: *“How many times have I dreamt at night that I was in this place, dressed, by the fire, although I was quite naked in my bed?”* (Descartes. R. 1968. p.96)

Cartesians believe this argument to show that although the body is in a dormant state, the mind is still active and that, although the brain may be active to some degree in a dreaming state, if we cut open the person’s head and looked inside we would only be able to observe the stimulated area; we would know nothing of what the person was actually dreaming and how they were perceiving it.

Objections to Dualism

Cartesian Dualism does not in any way assist us in understanding the nature of the mind; it only demonstrates that there is something is not physical that dreams, perceives, experiences and so on. Moreover, is the idea of a non-physical substance remotely coherent? – How might it be scientifically investigated, for example? And, more to the point, how can something non-physical interact in a causal way with a physical object?

We know that brain damage leads to mental deficiency - if the mind and body are separate how can this state of affairs be explained? If a person is brain damaged in ways that negate the possibility of them experiencing happiness, for instance, then no matter what they think, they will never be happy.

Behaviourism

Behaviourism rejects the existence of a mind altogether. The behaviourist considers the idea of an immaterial mind to be a fiction (the philosopher Gilbert Ryle famously described it as the “myth of the ghost in the machine”); mental states are merely a shorthand way of describing our behavioural qualities and propensities in certain ways. Consequently, a person is only angry when they exhibit the outward signs such as shouting or stamping their feet in particular kinds of ways. Behavioural adjectives are, as such, merely labels attached to certain kinds of human actions. Rather than solve the mind/body problem, behaviourists claim to have abolished it.

Objections to Behaviourism

It might seem as if Behaviourism does not allow a distinction to be made between someone genuinely in pain and someone pretending to be. However, this objection is not as powerful as one might think insofar as the behaviourist can point out that it is possible to tell the difference between fake and authentic behaviour by examining the behaviour which preceded and succeeded it. An actor, for example, might be able to fool someone who saw the pain behaviour in isolation but once they had realised that he was an actor and that prior to and after them seeing his pain behaviour he was laughing and joking, they would realise that he was dissimulating.

A more serious objection is that behaviourism fails to take into account what it actually feels like to be in a specific mental state; the actual experience does not figure at all. Moreover, behaviourism asserts that one learns about one’s own beliefs in the same way as one learns about the beliefs of others i.e. by observation of behaviour. However one does not need to make observations of our own behaviour in order to know that one is in pain – an observation is vulnerable to error but one cannot be wrong about whether one is in pain or not.

What of those individuals who are completely paralysed? According to the behaviourist they could not have any mental experience at all – since they could not exhibit any outward signs they would be unable to feel pain. Yet

evidence suggests that those who have been paralysed and have recovered to a certain extent are certainly capable of feeling pain, and in many cases they also have an intense mental life (“Locked-in” syndrome, for example).

Type Identity Theory

Type-Identity Theory is a physicalist theory of mind. This means that it reduces the entire mind to a physical entity; indeed, type-identity theorists consider the mind to be identical with the brain.

Type-Identity Theorists have argued that, in fact, all of experience is ultimately answerable to physical events and that a far better approach to understanding mental states would be to posit a form of brain-body theory of mind. Such a theory, in its crudest form, states that one can make a simple identification of one kind of phenomenon with another: a mental state – e.g. seeing or being in pain – is no more than a brain state; both are physical and both are ‘inner’. Mental states (or brain states, since the two are equivalent on this conception), are thus only inner in the sense that what occurs does so inside our skulls; they are not inner in any sense that puts knowledge of them, in principle, beyond our reach or the reach of others (unlike Cartesian Dualism). The basis for such an approach is an attempt to eradicate the problems thrown up by – firstly – the thought that behaviour is the only criterion we have for the legitimate ascriptions of psychological predicates when it has become apparent that our inner states seem to be the grounds for such behaviour and – secondly – that the link between inner phenomenal experiences and overt behaviour is non-logical in the sense that no definite connection between behaviour and mental states can be made because we cannot non-inductively identify phenomenal states in others the way we can (through introspection) in ourselves.

Objections to Type-Identity Theory

Consider the following example. There are two people gazing at the “Andromeda Galaxy”. Person X is a professional astronomer and person Y has only a fleeting interest in the heavens. They are both looking at, and thinking about, the galaxy and therefore (according to Smart and type-identity theorists in general) their mental states and brain states would be identical. However, surely the superior knowledge of person X would mean that he perceived the galaxy in a different way to person Y e.g. Person X considers the scientific attributes of the galaxy such as that it is 2 million light years away etc, while person Y merely appreciates its aesthetic qualities. Thus (assuming for the sake of argument that the same part of the brain is stimulated in both cases), their brain states would be the same, but (due to their different perceptions) surely their mental states would be different.

The thought must be in the same location as the brain state. However thoughts do not appear to have exact settings in this sense. Firstly, there is the phenomenon of neuroplasticity – that is, the same behaviour can correlate to two different brain states in the same person. Consider, for instance, someone who has suffered a brain injury and then re-learned certain skills (forms of behaviour) that are concomitant with activity in a different area of the brain. Secondly, other animals can experience pain but their brains are very different from ours; this means that the same phenomenon (i.e. the experience of pain) is related to very different physical realities in each case (octopus brains are very different from human brains but both human beings and octopodes feel pain).

Furthermore, Type-Identity Theory also fails to deal with the problems of consciousness i.e. the personal qualities of our own experiences. The observance of neurological function cannot, in any sense, reveal the quality of the experience that a person is having.

Functionalism

A comparatively recent theory succeeding Type-Identity Theory. Functionalism is causal in that it suggests mental states (inner states) initiate behaviour. Put another way, it relies on the notions of input from the external world through the afferent nerves >>processed by the brain (interaction) >> output in the form of behaviour.

Hilary Putnam (1926 – 2016) was a great exponent of Functionalism likening the brain to an automated machine. The idea of the brain and nervous system as a computer clarifies Functionalism well: The brain resembles the hardware of a computer; the software which enables the hardware to carry out actions are mental states / thoughts, in turn, the functions that the computer as a whole performs are the direct result of commands from an external operator – in the case of humans, this is the environment with which we interact.

Functionalism maintains the brain-body dualist approach insofar as it posits a causal link between the brain and body. It also and links in agreeably, though with some differences, with ideas in the philosophy of perception: the brain gathers information via the nervous system from our perceptual apparatus (eyes, ears, nose etc.) and processes it in such a way so as to provide us with perception and perceptual content.

Functionalists do not categorize our various psychological predicates with particular brain-states but they have no need to do so; all that they are arguing is that such mental adjectives are names that describe their functional roles not the quality (or otherwise) of the experiences that we may (or may not) feel when such predicates can be appropriately applied.

Thus, we can see that functionalism is compatible with both dualist and identity theories of mind. Inner states can be understood (and identified) as playing causal roles in relation to function and understood in brain-state terms.

Objections to Functionalism

If functionalism is accurate then it would seem to ascribe mental states to things that are not conscious, e.g. a university operates in a functionalist capacity in that it interacts with the external environment in certain ways as a result of inputs and outputs – however one cannot reasonably suggest that a university is conscious i.e. that it has purely subjective experiences in the form of *qualia*.

Turing Machines

The basic idea is as follows: an average human being engages in various linguistic activities with another human and a machine; each is isolated from the other. The machine has been designed to react in ways humans do. If the human judge cannot tell the difference between the machine and the human, the machine has passed the test and, as such, should be understood as a fully fledged artificial intelligence.

Bibliography

Block, N. 'Trouble with Functionalism' in *The Nature of Mind*. Ed. Rosenthal, D.M. OUP. Oxford. 1991.

Descartes, R. *Discourse on Method and The Meditations*. 1637. Penguin Classics. Translation: Sutcliffe.F.E. 1968. London.

Smart, J.C.C. 'Sensations & Brain Processes' in *The Nature of Mind*. Ed. Rosenthal, D.M. OUP. Oxford. 1991.