

## *SOME COMMENTS ON PROFESSOR SEARLE'S REITH LECTURES*

Dr U.T. Place

Editors' Note: This paper considers the contents of Prof. Searle's first four Reith lectures. It was prepared in response to a request for some form of commentary suitable for first year undergraduates, in particular.

In his first lecture Professor Searle presented a version of the view that is generally, but rather misleadingly, known as the mind-brain identity theory. This description is misleading in that none of the philosophers who have advocated such a view in recent years have ever wanted to claim that the mind and brain are one and the same entity or substance. The claim is that all mental events and, on some accounts, all mental states also, are the same events and states as some, as yet unidentified, events in or states of the brain. It has never been held by anyone, so far as I know, that all brain events and brain states are mental events or mental states, since there is evidently a great deal that goes on in our brains that is not reflected in our mental life.

In my view Professor Searle's version of the mind-brain identity theory is defective because, like most philosophers who have held this kind of view in recent years, people like David Armstrong, Donald Davidson and Dan Dennett, he fails to draw what I regard as the all-important distinction between a mental event in which something changes either at or over time and a mental state in which something persists unchanged over time, as well as the distinction between two kinds of mental event, mental processes or activities which involve continuous change over time and instantaneous mental events which occur at moments of time, but are not extended over time. In fact, the important contrast here is the contrast between mental processes and mental states. Instantaneous mental events, it turns out, invariably consist in the intersection between an antecedent mental process and a subsequent and consequent mental state. For example, the instantaneous mental event of deciding to do something consists in the change which occurs at the point when an antecedent mental process of wondering what to do gives way to the subsequent and consequent mental state of intending to do what one has thereby decided to do.

In my paper 'Is consciousness a brain process?' which was published in the *British Journal of Psychology* in 1956 and which is generally taken as the starting point of recent philosophical discussions of the mind-brain identity theory, I argued that it is a reasonable scientific hypothesis not to be dismissed by philosophical argument alone, that consciousness, understood as only part of what we talk about as the mind, is a process in the brain. I took the term "consciousness" to embrace mental processes and activities like having a sensation, experiencing something, looking, watching, listening, savouring, enjoying, trying, pondering, thinking out or about something, wondering, imagining, picturing and dreaming, in so far as these processes and activities do not consist in publicly observable behaviour, like the head and eye movements involved in looking or the talking audibly to oneself in which thinking frequently consists. I excluded mental states from the scope of this theory because Gilbert Ryle in his book 'the Concept of Mind' published in 1949, had shown, to my satisfaction at least, that mental states like knowing, believing, understanding, wanting, intending or being excited, angry, afraid, disgusted, sad, relieved, pleased, ambitious, vain and intelligent, all consist in a disposition, either a capacity, as in the case of knowing, understanding and being intelligent, or a propensity, as in other cases, to talk and behave in a variety of broadly specifiable ways, if and when the appropriate contingency arises. Thus to say that someone wants an apple is to say that they are likely either to ask for an apple or accept one when the opportunity arises. Likewise, to say that someone knows the time is to say that they are able to answer the question 'what is the time?' correctly and to act accordingly if they have an appointment to keep at the time in question.

It seemed to me then when I wrote my paper, and still seems to me, quite unnecessary and indeed mistaken to suppose that such mental states consist in a private inner state to which only their owner has access. When we ask someone if they want an apple or if they know the time, we are not asking them to introspect their inner state and tell us about aspects of their inner life which we have otherwise no access to: we are either, as in the wanting case, offering to supply the individual with an apple or, as in the knowing case, testing the individual's ability to answer the question correctly. It is quite unnecessary to suppose that wanting an apple or knowing the time is an inner state in order to explain why it is that you can't tell simply by inspection what someone wants or knows, in the way that it is necessary to suppose that pain, for example, is an inner process in order to account for the fact that you can't always tell simply by inspection whether or not someone is in pain.

I am not suggesting there isn't an inner state of the brain microstructure on which the propensity to ask for or accept apples or the ability to supply the right answer to the question 'what is the time?' depends. I fully accept that there must be such a state of the brain microstructure just as there has to be a state of the microstructure of glass which makes it flexible or of a magnet which gives its magnetic properties. What I am denying is that dispositional properties consist in or are identical with the state of the microstructure of the entity to which the property belongs with which they are correlated. In other words, unlike the case of mental processes, which, except in so far as they involve adjustments of the sense organs and overt speech movements, I want to say simply *are* processes in the brain, mental states and the brain states with which they are correlated are two distinct and causally related things. Mental states are dispositions to talk and behave in a variety of publicly observable ways which depend for their existence on, but are not identical with, certain as yet undiscovered states of the brain microstructure.

Searle in his first lecture is obviously pulled both ways on this issue. He sees the force of the contention that mental states are causally dependent on the state of the brain microstructure with which they are presently correlated. At the same time he is attracted by the view first put forward by David Armstrong in his book 'Materialist Theory of the Mind' (1968) in which he suggested that the mind-brain identity theory needs to be extended from the case of consciousness and mental processes to which I had restricted it in my paper to the case of mental states, thus yielding a uniform account of mind as a whole. These conflicting inclinations lead Searle to adopt the extraordinary, not to say absurd, view that mental states are both identical with and causally dependent on the state of the brain microstructure with which they are correlated. Searle is here trying to have his cake and eat it which, as we all know, can't be done. Either mental states and the corresponding states of the brain are two descriptions of one and the same state or they are two distinct and causally related states. You can't have it both ways, as Searle thinks he can.

Turning to Searle's second lecture, we now begin to see some of the disastrous consequences which follow from Armstrong's attempt to stuff mental states into the brain where, on a proper understanding of them, they simply do not belong. For the upshot of Searle's Chinese room argument to which most of the second lecture was devoted is not, as he seems to think, that the brain is not a computer, but rather a biological system endowed with the mysterious mental property of intentionality. It is simply that a mental state like that of understanding an utterance is just not, as Searle takes it to be, a state of the brain. It is a matter, in the case of a command or request, of being able, given the relevant physical capacities, to do what one has been asked to do, in the case of a question, of being able, given the relevant information, to supply the appropriate answer and, in the case of a statement, of being able to make effective use, given that the statement is accepted as true, of the information it contains. In other words understanding what someone says presupposes a community of agents each of which is pursuing its own individual set of

objectives within a common environment and communicating with other agents by means of a language whose intelligibility depends on syntactic and semantic conventions which are established and maintained by the verbal community constituted by speakers of the particular natural language in question. Computers, if they are programmed to do so, do have a limited ability to obey commands, answer questions and act on information received in the form of an indicative sentence in a natural language such as English. What a present day computer does not have is the ability to interpret and construct indefinitely many sentences which it has never previously encountered or constructed and relate these sentences to features of a complex physical and social environment as they are detected by the kind of sensory equipment available to the average human being.

A second mistake which appears in Searle's second lecture concerns the account he gives of the mysterious mental property known as "intentionality" which is supposed to distinguish mental states and processes from non-mental or physical states or processes.

Intentionality, in the sense in which Searle discusses it in his second lecture and at greater length in a recent book, is a notion which derives from the work of the 19th century German philosopher and psychologist Franz Brentano. According to Brentano, intentionality is that feature of mental states and processes whereby they are orientated towards an object which, either as in the case of a goal or objective, does not yet exist and may never do so or, as in the case of an object of thought, may exist, but may equally well be a non-existent person such as Father Christmas or a non-existent state of affairs such as Utopia. However thanks to Dr John Burnheim of the Department of General Philosophy, University of Sydney, who made the point in an unpublished paper presented to the then united Department of Philosophy at Sydney some fifteen years ago, we now know that intentionality, so far from being, as Brentano thought, the mark of the mental, is the mark of the dispositional.

Intentionality, as thanks to Burnheim we now understand it, is the teleological feature of dispositions whereby they involve an orientation of the entity which possesses the disposition towards the manifestation of a particular kind of behaviour, if at any time certain conditions are fulfilled. Thus the brittleness of a pane of glass is intentional in so far as it involves an orientation of the glass towards the occurrence of an event, its breaking, which has not yet occurred, may never occur (at the end of the day it may just melt in a fire) but which may occur at any time, if the appropriate conditions are fulfilled. Clearly since mental states all consist of some kind of dispositions, they are intentional and so are some mental activities like looking for something, enjoying something or trying to do something, all of which involve some kind of mental disposition with which the mental activity is performed, but since there are mental activities like looking at something or listening to something which need not involve any dispositions, we have to accept that there are non-intentional mental processes as well as many non-mental dispositions and therefore many non-mental intentional states like brittleness.

Professor Searle's third lecture was devoted to a critique of Cognitive Science. His central argument, the argument that cognitive science confuses two different senses of information and information processing is one which in my view is absolutely right. Unfortunately his characterisation of these two senses of "information" is somewhat defective. One sense of information is the ordinary sense in which we talk about one person conveying information to another either by spoken or written language or by some other form of sign, gesture or combination of signs and gestures. In this sense of "information" it always makes sense to ask of a bit or piece of information 'What is it information about?' and to express the information conveyed in the form of one or more sentences in some natural language. The other sense of

'information' is the sense in which that term is used in the context of what is known as 'information theory'. Information theory is a precisely defined mathematical theory which was developed in the first instance by and for engineers concerned with the problems of telecommunications - the transmission of information in the form of electrical impulses along a telephone wire, in the form of radio waves through space or through the atmosphere, in the form of a laser beam, light along an optic fibre etc. For the purposes of this technology the actual content of the message, what it means, what it is about, is irrelevant. Consequently a bit of information is defined within information theory in such a way that it does not make sense to ask of a bit of information, so defined 'what is this bit of information about?' Now while it is true that the important difference between these two senses of 'information' is ignored by cognitive science, so-called, to describe the information theory sense of 'information', in the way that Searle does, as 'as-if information' is grotesque when you appreciate the very precise technical meaning that is given to the term by its definition within information theory. Searle then goes on to compound his felony by talking about 'as-is information processing' when the concept of 'information processing' has its original and natural home as a technical concept within information theory. Any use of the expression to refer to an operation carried on with information in the ordinary sense of the word is a metaphor from the technical usage of information processing which is itself a metaphor derived from the ordinary use of 'information'.

In addition to the distinction he draws between 'genuine' and 'as-if information processing', Searle also draws a distinction between 'genuine rule-following' which involves understanding the rule that is being followed and 'as-if rule-following' which does not. He suggests that the rule following that applies in the case of computers is of the 'as-if' kind. I would not want to dispute either the distinction or his application of it to the case of computers. What I would dispute is his claim that linguistic rules are rules of the genuine kind rather than behaviour which is as if a rule was being followed. It seems to me that, in the sense in which we can speak of genuine non-metaphorical rule-following, a rule is a universal imperative sentence, like '*Always close the door when you come into the room*', and to follow a rule is to understand such a sentence and be guided by it. We *do* follow rules in this sense when we learn certain linguistic skills like spelling and punctuation and again when we learn the grammar and semantics of a second language in the way that such languages are normally taught in school. But we do not and could not learn a first language in this way, for the simple reason that we would have to understand the language in which the rule was formulated before we could begin to learn the language itself, which is absurd. The rule-following involved in the interpretation of sentences by native speakers of a language is as-if rule-following, not genuine rule-following.

Since I first gave this talk last week, we have heard Searle's fourth lecture on the theory on Action. This lecture actually contained so many misconceptions that it is difficult to know where to begin. I must say that I found it peculiarly irritating to have to listen to a collection of tired cliches that have been around in the philosophy of action since the early 1960's, and which I have been warning my philosophy of mind students to avoid ever since. I first joined this department fifteen years ago, all trotted out with massive self-confidence as if some new and brilliantly original theory were being revealed for the first time.

All I can do, I think, is simply to list some of the mistakes Searle makes in this lecture without attempting to explain why these *are* mistaken. Searle began his lecture by trying a distinction between two kinds of behaviour which he called 'movement' and 'action' and made the demonstrably false claim that behaviourism can only account for movements and not for actions. He then committed himself to the mistaken view that all actions - in other words anything that anyone can be said to do - are characterised in terms of the intention with which they are done and

from this he drew the conclusion, which is true only of one understanding of what it means to act on an intention, that the only kind of explanation that can appropriately be given for an action so defined is an explanation in terms of what the agent wants to achieve or avoid and what he or she knows, believes or thinks about the way in which that goal or objective can be achieved. Searle's misunderstanding of the notion of intentionality was again apparent in this lecture, as was his gross misunderstanding of what is involved in a causal relation. This weakness of his theory of causation appears both in his failure to appreciate the relation between causation and action in a sense in which one person or thing which we may call the agent acts on or does something to another person or thing, the patient, and in his notion of intentional causation where he seemed to be creating an aura of mystery around the observation that an intentional cause - by which I take him to mean an intention to do something - involves a representation of the effect it is intended to bring about.

In this connection Searle *did* introduce an important distinction between two senses of acting from an intention. The contrast here was between spontaneous actions which are done on the spur of the moment and actions which are carried out as the result of an earlier decision and subsequent intention to carry out the action as and when the appropriate opportunity arises. The point here is that even though I may be acting on impulse when I put my arm around someone, there is an important sense in which that action is intentional and involves a representation of its intended effect. But this intention and this representation of the intended effect is very different in kind from the intention and representation of the action one intends to do in the case of a premeditated action. In the premeditated case the intention to do something is formed as a result of a piece of verbally formulated reasoning. In this case the representation of the action to be performed takes the form of the sentence which if not grammatically imperative nevertheless has the form of a self directed imperative by means of which the agent resolves to do what he or she has decided to do. What Searle fails to appreciate is that only actions based on this kind of verbally premeditated intention involve practical reasoning of the kind he describes and which is pre-supposed by the use of the knowledge, belief, thought, desire type of explanation. We do, of course, use this kind of explanation to account for the behaviour of animals and human infants who simply do not have the verbal ability to work out a pre-conceived plan of action in verbal terms, but also for a great many actions of human adults who *do* have the ability to do this, but do not exercise it when they act spontaneously and impulsively.

The sense in which I can be said to intend to put my arm around someone when, acting on impulse, I try to do so, is to be understood in my view in terms of the cybernetic notion of negative feedback. In order to me to move my arm in such a way that it ends up around my neighbour's shoulder, I have to pay attention to the visual and kinaesthetic feedback from the movement as it develops and make the necessary corrections as and when it tends either in the direction of falling short or in the direction of overshooting. But this, as you will see, implies some kind of preconception of how the movement is to turn out of what the end result is to be. This preconception, however, is not verbally formulated, nor does it consist in some kind of visual or kinaesthetic image of the state of affairs to be achieved at the end of the day, since you would need the preconception to be able to form the mental image in the first place. The best way of thinking of it, I suggest, is in terms of a thermostatically controlled central heating system which also operates on the principle of negative feedback. In terms of this model the preconception involved in this kind of intention action is like the thermostat setting the central heating system - very mysterious if you try to describe it as if it were an inner mental state to be discovered by introspection - very simple when you think of it as dispositional property of a negative feedback system involving, not just the brain, but two external feedback loops, one through the muscles, the other involving light rays reflected back onto the retina of the eye from the movement as it develops.