

Contemporary Reactions to Descartes' Philosophy of Mind

Quassim Cassam

Overview

It is widely assumed that Descartes' philosophy of mind is organized around three major commitments. The first is to substance dualism. The second is to individualism about mental content. The third is to a particularly strong form of the doctrine of privileged first-person access. Each of these commitments has been questioned by contemporary philosophers of mind. Substance dualism is generally regarded as a non-starter, individualism has come under attack from a number of different quarters, and the doctrine of privileged access has been watered down or rejected. Yet, at least as far as questions about mental content and privileged access are concerned, contemporary discussions still address what they represent as Descartes' views. More often than not crude parodies of these views end up as the focus of discussion but more careful critics are usually prepared to recognize that Descartes' philosophy of mind is more subtle and nuanced than the parodies might lead one to suppose.

Responses to substance dualism, the view that mind and body are distinct substances one of which (body) is material and the other (mind) immaterial, fall into two main categories. There are those which question its coherence and those which reject it on empirical grounds. It remains to be seen which form of objection is more appropriate but it is worth noting that some critics of substance dualism have been prepared to endorse another kind of dualism, a dualism of properties. According to this 'dual aspect' version of dualism mental properties are neither identical with nor reducible to physical properties, even though both mental and physical properties are properties or aspects of a single substance. This wouldn't have satisfied Descartes but it may well be the best that can be done for dualism in the philosophy of mind.

Individualism is roughly the view that which thoughts a person can have does not depend on his or her relations to the physical or social environment. A person's thoughts are, in this sense, 'world-independent'. Individualism about the mental, also known as 'internalism', is often attributed to Descartes on the basis of a reading of his thought experiments in the First Meditation. On this reading, Descartes is committed to individualism because he envisages the possibility of our being radically mistaken about the nature and existence of the world and of our thoughts remaining just as they are in these circumstances. In response, it has been claimed that it is a mistake to move from the premise that our thoughts about the world could be radically mistaken to the conclusion that they are individuated individualistically and that there are in any case good independent arguments against individualism. From an anti-individualist perspective, therefore, Descartes' conception of mental content is of interest because it brings the defects of individualism into the sharpest possible focus.

The doctrine of privileged first-person access says that one's introspectively based judgements about one's own mental states enjoy a range of epistemic privileges that judgements about non-mental reality or the mental states of others do not enjoy. One of these privileges is infallibility or immunity to error. Immunity to error does not entail immunity to ignorance but the strongest versions of the doctrine of privileged access insist on both forms of immunity. They claim that introspectively based judgements about one's own mental states can't be mistaken and that one can't fail to know what is in one's own mind. On the face of it both of these theses are too strong. Yet despite the fact that neither ignorance nor error can be ruled out with respect to many states of mind there does nevertheless appear to be something right about the doctrine of privileged access. For example, one might think that the basis on which one ascribes thinking to oneself is different from the basis on which one ascribes it to others and that at least some of one's

judgements about one's own mind can't be mistaken. On this account the challenge is to explain the authority of self-knowledge without exaggerating its strength or scope.

As well as raising questions about the mind-body relation, mental content and privileged access Descartes' philosophy of mind raises questions about the relationship between these issues. Some materialist critics of dualism have argued that the doctrine of privileged access implies the falsity of materialism, and that arguments for materialism and against dualism are therefore also implicitly arguments against privileged access. Other commentators have represented Descartes as arguing for individualism on the basis that one's judgements about own thoughts are infallible. This has in turn sparked a debate between those who have been prepared to concede that one's epistemic access to one's own thoughts can't be privileged in this way unless individualism is true and others who have argued that we can know our own thoughts in an authoritative manner even if individualism is false.

These are complex issues and matters are further complicated by questions about the strength of Descartes' commitment to dualism, individualism and the doctrine of privileged access. One suggestion is that Descartes' position on the mind-body relation is actually a form of 'trialism', according to which thoughts are assignable to mind, extension to body and sensations to the union of mind and body. Others have drawn on the Third Meditation to argue that the attribution of individualism to Descartes is not well grounded. It is also possible to find commentators who read Descartes as holding that introspective judgements are only privileged up to a point and that such judgements are not absolutely immune to error. These questions of interpretation remain unresolved. Nevertheless, it is worth bearing them in mind since they open up the possibility that contemporary responses to dualism, individualism and the doctrine of privileged access are not necessarily

responses to views which Descartes actually held. For the moment, however, let us ignore such interpretive worries and take a closer look at these responses themselves.

Dualism

P. F. Strawson writes somewhere that one of the marks of a really great philosopher is to have made a really great mistake. He goes on to argue that Cartesian dualism is one such mistake, its greatness consisting in the fact that it gives a ‘persuasive and lastingly influential form to one of those fundamental misconceptions to which the human intellect is prone when it concerns itself with the ultimate categories of thought’ (Strawson 1974: 169). But why is Cartesian dualism a mistake? According to Strawson and many others the fundamental problem is that this form of dualism is not just false but incoherent. For the notion of an immaterial Cartesian mind or soul to make sense it must be possible for specify criteria of singularity and identity for souls. That is to say, ‘we must know the difference between one such item and two’ and ‘we must know how to identify the same item at different times’ (Strawson 1974: 173). Since bodies are in space as well as time we can account for their singularity and identity in spatio-temporal terms. For example, we can appeal to the principle that two bodies can’t occupy exactly the same region of space at the same time. But the fact that immaterial souls are supposed to be non-spatial leaves us without any conception of what their singularity and identity consists in. That is why, according to the present line of thinking, Cartesian dualism is conceptually incoherent.

One response to this objection would be to argue that it is possible to count and reidentify souls by reference to the human beings or human bodies to which they are attached. Where there is one human being we assume that there is, or was, one soul attached to it and that sameness of human being implies sameness of soul. Yet it is not clear how this assumption can be justified. Strawson makes this point by means of the following example:

Suppose that I were in a debate with a Cartesian philosopher, say Professor X. If I were to suggest when the man, Professor X speaks, there are a thousand souls simultaneously thinking the thoughts his words express, having qualitatively indistinguishable experiences such as he, the man, would claim, how would he persuade me that there was only one such soul? (1974: 174).

On the face of it, a substance dualist needn't be troubled by this question. He might not be able to persuade an outside observer of this but he may nevertheless claim to be directly acquainted with the singularity and identity of his thinking self. In effect, this amounts to the suggestion that one can be conscious that one's thoughts belong to one and the same immaterial soul even if one is unable to give any informative general account of the criteria of singularity and identity for souls.

But is the identity of one's own soul something with which one can be acquainted in this way? How can I rule out the possibility that what I am conscious of as one persisting soul is in fact a series of distinct souls each of which transmits its states of consciousness to its immediate successor? In response to these questions the substance dualist ought to argue that the simplest and best explanation of the evident unity of one's consciousness is that one's mental life is underpinned by a single soul rather than a succession of souls. To claim that this is the best explanation is not to claim that it can't be mistaken but the fact that one can't completely rule out the 'multiple souls' hypothesis doesn't show that it doesn't make sense to talk about the singularity and identity of souls. The most that can be said is that the dualist who claims consciousness of the singularity and identity of his immaterial soul faces an epistemological problem but Strawson's point was supposed to be conceptual rather than merely epistemological.

It is easy to overlook the distinction between epistemological and conceptual considerations if, like Strawson, one subscribes to the verificationist principle that 'you do

not know what souls are unless you know how to tell one from another and to say when you have the same again' (1997: 51). Given this principle, it is tempting to argue for the incoherence of substance dualism on the basis that one can't tell one soul from another on a suitably strong reading of 'tell'. Yet, as we have seen, the singularity of souls is something for which the unity of consciousness might be seen as providing at least defeasible evidence and it is not clear in any case why one should accept Strawson's verificationism. So we still lack a conclusive demonstration of the incoherence of Cartesian dualism.

This has prompted some philosophers of mind to pursue a different line of argument against dualism. For example, Parfit explicitly rejects the suggestion that the concept of a Cartesian soul or ego is unintelligible and concedes that there might have been evidence supporting the Cartesian view. Specifically, if there was sufficient evidence for reincarnation we might reasonably conclude that a Cartesian ego is what each of us really is. The problem, according to Parfit, is that we lack good evidence for the belief in reincarnation. Hence, 'even if we can understand the concept of a Cartesian Pure Ego, or spiritual substance, we do not have evidence to believe that such entities exist' (1984: 228). Unlike Strawson, therefore, Parfit and others like him reject Descartes' dualism on empirical rather than conceptual grounds.

It remains controversial whether the rejection of Cartesian dualism on empirical grounds is warranted or whether it is appropriate to criticize Descartes' position on anything other than conceptual grounds. What is clear is that few contemporary philosophers of mind regard substance dualism as a serious option. They assume that there are decisive objections to it but they often fail to spell out these objections in any detail. Parfit's approach has not gained widespread acceptance and it continues to be assumed, often without much argument, that substance dualism makes no sense. While this might ultimately be the right thing to think we have seen that the charge of incoherence is less

easy to justify than one might initially have supposed. Faced with the objection that substance dualism can't account for the singularity and identity of souls there are several points at which the dualist can dig in his heels, and the same goes for other standard arguments for the incoherence of dualism. If this is right then substance dualism has been rejected rather than refuted.

Nevertheless, it is easy to see why substance dualism looks so much less attractive to us than to it did to Descartes. If we think of the world as the natural world and as causally closed there won't be any room in it for a 'separate realm of mental substance that exerts its own influence on physical processes' (Chalmers 1996: 124-5). If there is a sufficient physical cause for every physical event 'there is no room for a mental "ghost in the machine" to do any extra causal work' (Chalmers 1996: 125). What, in that case, would a naturalistic conception of mind look like? A reductive naturalism would not only make no room for a separate realm of mental substances but also regard mental properties like pains as identical with physical properties like C-fibre firings. This is the view of type-type identity theorists such as U. T. Place, J. J. C. Smart and D. M. Armstrong. Yet this form of reductive naturalism has been widely criticized and one cannot fail to be struck by the Cartesian overtones of some of the best known criticisms.

Consider Saul Kripke's anti-materialist argument in Naming and Necessity. Kripke argues that 'the identity of pain with the stimulation of C-fibers, if true, must be necessary' (1980: 149). At any rate, this is what one would expect if such type-type identities are analogous with such scientific type-type identifications as the identity of heat with molecular motion. Yet the possibility, or apparent possibility, of C-fibre stimulation without pain and pain without C-fibre stimulation suggests that the correspondence between the two has 'a certain obvious element of contingency' (1980: 154). This apparent contingency cannot be explained away in the way that the apparent contingency of the

correlation between heat and molecular motion can be explained away. In the latter case there is a distinction to be drawn between heat and the sensation of heat so that the apparent possibility of molecular motion without heat is really only the possibility of molecular motion without the sensation of heat. But to conceive of C-fibre stimulation without the sensation of pain is to conceive of C-fibre stimulation with pain; there is no distinction in this case between pain and the sensation of pain. Thus, when God created the world all he needed to do to create heat was to create molecular motion. But when he created C-fibre stimulations he still had more work to do to create pain, that is, to ensure that C-fibre stimulations are felt as pain.

Like Descartes, Kripke takes it that a certain kind of conceivability is a guide to possibility. Descartes argues for a real distinction between mind and body on the basis that they can be understood or conceived of apart from one another. Analogously, Kripke argues for the non-identity of pain and C-fibre stimulation on the basis that each can be conceived of as existing without the other. Kripke's intuitions are, in this sense, Cartesian and he relies on his Cartesian intuitions to argue for the falsity of some types of materialism. Yet Kripke is no Cartesian dualist. In his view, a person could not have come from a different sperm and egg from the ones from which he actually originated. This 'implicitly suggests a rejection of the Cartesian picture' (1980: 155 n.77) because there is no obvious reason why an immaterial soul should have any necessary connection with a particular sperm or particular egg.

One alternative to Cartesian dualism is what Chalmers calls 'naturalistic dualism'. Chalmers argues consciousness is not logically supervenient on the physical in the sense that 'all the microphysical facts in the world do not entail the facts about consciousness' (1996: 93). One argument for this claim appeals to the logical possibility of zombies. My zombie twin is molecule for molecule identical with me but lacks conscious experience

entirely. If such a zombie is conceivable that is enough to establish that consciousness cannot be reductively explained. Yet it doesn't follow from the fact that consciousness doesn't supervene logically on the physical that it doesn't supervene naturally on the physical. The naturalistic dualism for which Chalmers argues is a form of property dualism. The idea is that 'conscious experience involves properties of an individual that are not entailed by the physical properties of that individual, although they may depend lawfully on those properties' (1996: 125). What makes this a form of naturalism is its insistence that 'we can explain consciousness in terms of basic natural laws' (1996: 128). Accordingly, consciousness turns out to be just another natural phenomenon even though conscious properties can't be reduced to physical properties.

The limited concessions to dualism that writers such as Kripke and Chalmers are prepared to make have to do with their conception of the nature of consciousness and of the relation between conscious properties and physical properties. Yet there are also intentional states which lack any distinctive conscious character. What makes a pain a pain is the way it feels to its subject whereas what makes the belief that George W. Bush is the American President the belief it is isn't the way it feels to believe that George W. Bush is the American President. There isn't anything specific that it is like to have this belief in the way that there is something it is like to be in pain. Rather the belief that George W. Bush is the American President is the belief that it is partly in virtue of its logical or normative relations to other beliefs. Some have seen this as undermining against any attempt to identify beliefs with physical or functional properties. Yet this doesn't count in favour of Cartesian dualism any more than the failure of materialist explanations of consciousness counts in favour of Cartesian dualism. The appeal to a non-physical mind is of no help in either case since it is quite obscure how it is any easier to explain consciousness or intentionality in non-physical terms than it is to explain them in physical terms. What

contemporary philosophers have extracted from Descartes' philosophy of mind is therefore not a solution to the mind-body problem but a sense of its depth and intractability. The last word goes to Kripke, who concludes his discussion with the observation that the mind-body problem is 'wide open and extremely confusing' (1980: 155 n.77).

Individualism

Individualism can roughly be characterized as the view that 'one's mental phenomena are in certain fundamental ways independent of the nature of the empirical and social worlds' (Burge 1986: 120). More precisely, it is the view that:

an individual person's or animal's mental state and event kinds.... can in principle be individuated in complete independence of the natures of empirical objects, properties, or relations (excepting those in the individual's own body, on materialist and functionalist views) – and similarly do not depend essentially on the natures of the minds or activities of other (non-divine) individuals (Burge 1986: 118-9).

According to Burge, 'individualism as a theory of mind derives from Descartes' (1986: 117). Specifically, it derives from a particular reading of Descartes' thought experiments in the First Meditation. What these thought experiments show is that our beliefs about what the empirical world is like could be radically mistaken, and this might lead one to conclude that the individuation of thoughts is unaffected by possible differences in the environment. Yet, Burge argues, no such conclusion is warranted by Descartes' thoughts experiments. We can concede that our thoughts about the world might be radically mistaken without conceding anything about 'how our thoughts about the world are determined to be what they are' (Burge 1986: 122).

If it is true that Descartes is committed to individualism then arguments against individualism are also arguments against Descartes' philosophy of mind. Many such arguments have been proposed. A key paper in this connection is Hilary Putnam's 'The

Meaning of “Meaning”. Putnam imagines a planet called Twin Earth which is exactly like Earth except that the liquid called “water” on Twin Earth isn’t H₂O but a different liquid whose long chemical formula we can abbreviate as ‘XYZ’. XYZ is indistinguishable from water at normal temperatures and pressures, and the oceans and lakes of Twin Earth contain XYZ rather than water. Now imagine a time when a typical Earthian speaker of English – call him Oscar₁ - didn’t know that water is H₂O and when the typical Twin Earthian speaker of English – call him Oscar₂ - didn’t know that “water” is XYZ. Even if the two Oscars were ‘exact duplicates in appearance, feelings, thoughts and interior monologue etc’ (Putnam 1975: 224) the extension of the term Earthian term “water” – the stuff that the term was true of – was still H₂O at this time and the extension of the Twin Earthian term “water” was still XYZ. In this sense the two Oscars understood the term “water” differently even though they were in the same psychological state. Putnam’s conclusion is that ‘the extension of the term “water” (and, in fact, its meaning in the intuitive preanalytical usage of that term) is not a function of the psychological state of the speaker by itself’ (1975: 224).

This is not yet an argument against individualism. Indeed, the suggestion that Oscar₁ and Oscar₂ are exact duplicates in feelings, thoughts and interior monologue despite the difference in their physical environments looks like an argument for individualism rather than an argument against it. Yet there is an obvious objection to Putnam’s own reading of his example. The objection is that Oscar₂ couldn’t possibly be thinking thoughts involving the concept water since he has never had any contact with water or with anyone else who has had contact with water. Since Oscar₁ has been in contact with water and does employ the concept water in some of his thoughts this is at least one respect in which their thoughts must be different. However, as Burge points out, this difference in their thoughts, in their mental states, derives from differences in their environments. This is now an

argument against individualism since this kind of dependence of the two Oscars' mental phenomena on their physical environments is precisely what individualism is committed to denying.

Other Burgean arguments emphasize the way in which one's mental states depend essentially on the nature of one's social environment. Suppose, for example, that a patient has the false belief that he has developed arthritis in his thigh. This must be a false belief since arthritis is specifically an inflammation of joints. But now imagine a counterfactual situation in which the patient's physical, behavioural and dispositional history is exactly the same as in the actual world but in which the word "arthritis" is conventionally applied to various rheumatoid ailments as well as to arthritis. In this counterfactual situation the patient lacks the belief that he has arthritis in the thigh. He couldn't have picked up the concept arthritis because "arthritis" in the counterfactual community doesn't mean arthritis. The upshot is that 'the patient's mental contents differ while his entire physical and non-intentional mental histories, considered in isolation from their social context remain the same' (Burge 1998: 28). This difference is attributable to differences in his social environment just as, in other cases, differences in mental content are attributable to differences in the physical environment.

If successful, these arguments against individualism are also arguments against what Putnam calls 'the assumption of methodological solipsism'. This is the assumption that 'no psychological state, properly so-called, presupposes the existence of any individual other than the subject to whom that state is ascribed' (1975: 220). Putnam claims that this assumption is 'pretty explicit in Descartes' (ibid.) and Fodor is making the same point when he attributes to Descartes the view that 'there is an important sense in which how the world is makes no difference to one's mental states' (1975: 228). But if anti-individualism is correct then how the world is does make a difference to one's mental states and there are

psychological states which presuppose the existence of individuals other than the subject to whom the states are ascribed. The thoughts that one can think are constrained by the concepts that are available to one and the concepts that are available to one are not independent of one's physical and social environment.

Why, in that case, should contemporary philosophers of mind ever have been attracted by Descartes' alleged individualism or methodological solipsism? One explanation is that this approach is in keeping with the idea that mental processes are computational. As Fodor puts it:

Insofar as we think of mental processes as computational (...) it will be natural to take the mind to be, inter alia, a kind of computer.... If we want to extend the computational metaphor by providing access to information about the environment, we can think of the computer as having access to "oracles" which serve, on occasion, to enter information in the memory.... The point is that, so long as we are thinking of mental processes as purely computational, the bearing of environmental information upon such processes is exhausted by the formal character of whatever the oracles write on the tape. In particular, it doesn't matter to such processes whether what the oracles write is true; whether, for example, they really are transducers faithfully mirroring the state of the environment, or merely the output end of a typewriter manipulated by a Cartesian demon bent on deceiving the machine (1975: 230-1).

On this account, the computational picture of the mind makes sense of 'the Cartesian claim that the character of mental processes is somehow independent of their environmental causes and effects' (Fodor 1975: 231). It also purports to provide the best explanation of the subject's behaviour. It is what an agent has in mind – his beliefs and desires, for example- that causes his behaviour. So if it turns out that the computational picture of what

the agent has in mind is best placed to explain what the agent does then that will be a powerful argument for the methodological solipsism to which this picture is committed.

It is controversial, to say the least, whether this argument for methodological solipsism is successful. In particular, it has been objected that the behavioural argument for methodological solipsism fails because it relies on an impoverished conception of 'behaviour'. According to this line of thinking, explaining what an agent does is not just a matter of explaining a series of bodily movements or motor responses. What is required is an explanation of the agent's actions, and this means that 'one cannot leave the truth or falsity of agents' beliefs out of account' (Hornsby 1986: 107). In this sense, it does matter whether what the oracles write is true. It matters because 'a person can be expected to do what she tries to do on occasion only if certain beliefs that explain her then trying to do that are true' (ibid.).

Be that as it may, methodological solipsists have other argumentative resources at their disposal. Perhaps the most influential argument for methodological solipsism or individualism is that we have privileged epistemic access to the contents of our own minds and that only individualism can explain how this is possible. For suppose that one's mental phenomena are, as some anti-individualists maintain, dependent on the nature of one's physical environment. In that case, given that one can be mistaken about the nature of one's physical environment, it would seem to follow straightforwardly that one can also be mistaken about one's own mental phenomena. So if one thinks that one can't be mistaken about the contents of one's own mind then individualism about the mental looks like the only serious option.

To assess this argument we will need to take a closer look at the doctrine of privileged access. Before doing that there is an important historical question that needs to be addressed. The question is whether it is correct to read Descartes as an individualist.

Consider his argument for the existence of God in the Third Meditation. Descartes argues for God's existence on the basis that God must be the source his idea of God. The implication is that idea of God depends in a fundamental way on the thinker's being embedded in a particular 'cosmic' environment. And if the idea of God is, in this sense, 'world-dependent', then so are those mental contents in which this idea is deployed. It is therefore false, even by Descartes' own lights, that how the world is makes no difference to one's mental states.

In fact, this attempt to read Descartes as a proto-anti-individualist or 'externalist' is too quick. The dependence that anti-individualism is interested in is the dependence of one's mental states on the nature of the empirical and social worlds. Since God is not a constituent of empirical or social reality the dependence of the idea of God on God's existence does not count against individualism. Indeed, it is interesting to note that Burge's characterization of individualism explicitly addresses this issue. If, as he stipulates, individualism is the view that an individual person's or animal's mental state and event kinds do not depend essentially on the natures of the minds or activities of other non-divine individuals, the fact that there are ideas which depend on the nature or activities of God doesn't look like placing Descartes in the anti-individualist camp.

Nevertheless, there is something right about the thought that there are elements of anti-individualism in Descartes' thinking. Indeed, Burge concedes that his earlier attribution of individualism to Descartes was 'badly grounded' (2003: 291). The principle to which Descartes appeals in the Third Meditation is that the cause of an idea must have at least as such formal reality as there is objective reality or intentional content in the idea, and this principle 'seems distinctly anti-individualist in spirit' (Burge 2003: 293). This brings us back to relationship between anti-individualism and the doctrine of privileged access. If it is true that Descartes subscribes to this doctrine, and that only individualism

can make sense of the ways in which self-knowledge is epistemically privileged, then it might seem uncharitable to try to read Descartes as an anti-individualist. In practice, however, anti-individualists tend to argue that anti-individualism is compatible with respectable versions of the doctrine of privileged access. It is now time to consider whether they are right about this.

Privileged Access

The doctrine of privileged access is one element of what Ryle calls ‘Descartes’ Myth’. According to this doctrine:

a person has direct knowledge of the best imaginable kind of the workings of his own mind. Mental states and processes are (or are normally) conscious states and processes, and the consciousness which irradiates them can engender no illusions and leaves the door open for no doubts. A person’s present thinkings, feelings and willings, his perceivings, rememberings and imaginings are intrinsically “phosphorescent”; their existence and their nature are inevitably betrayed to their owner (Ryle 1949: 15).

To say that a person has knowledge of the best imaginable kind of the workings of his own mind is, among other things, to say that his introspectively based beliefs about his own mental states and processes can’t be mistaken. In other words, such beliefs are infallible. This is one dimension of the doctrine of privileged access. A different dimension is at issue in the suggestion that one’s own mental states and processes are “phosphorescent”. To rule out the possibility of ignorance with regard to the nature and existence of one’s mental states and processes is to regard such states and processes as inherently self-intimating. What this means is that it isn’t possible for a proposition ascribing current mental states or processes to oneself to be true without one’s knowing that it is true.

How does the doctrine of privileged access relate to Cartesian dualism? On the one hand, one might think that conceiving of the mind as an immaterial spiritual substance does not commit one to regarding its activities as perfectly transparent to itself in the way that the doctrine of privileged access implies. On the other hand, some materialists have argued this doctrine is incompatible with their conception of the nature of mind. For example, Armstrong defends a version of Central-state Materialism according to which mental processes are states of the person apt for the production of certain sorts of behaviour. Yet knowledge of causes cannot be infallible or, as Armstrong puts it, 'incurable'. Accordingly, 'it is essential.... for the defender of Central-state Materialism to show that there can be no logically indubitable knowledge of, or logically privileged access to, or self-intimation by, our current mental states' (Armstrong 1968: 103). If it turns out that those states of a person that are apt for the production of certain sorts of behaviour are in fact physical states of the brain then introspection will have to be a physical process in the brain.

This explains why some materialists have been opposed to the doctrine of privileged access but it does not explain what is wrong with this doctrine. Objections to infallibility and self-intimation can be more or less radical. Less radical critics of the doctrine of privileged access concede that this doctrine might apply to a restricted class of mental events, namely sensations. So, for example, Boghossian remarks that 'it seems not conceivable, in respect of facts about pain, that we should be either ignorant of their existence or mistaken about their character, just as the Cartesian doctrine requires' (1998: 151). On this account it is in respect of thoughts and emotions that Descartes goes wrong. We can be both mistaken and ignorant about our own thoughts and emotions so only a restricted version of the doctrine of privileged access has any chance of being acceptable. In contrast, radical critics of this doctrine maintain that it isn't even true of sensations like

pain. One can be in pain without realizing it and one can think that one is in pain when one isn't.

In these terms, Armstrong is an example of a radical critic of privileged access. He gives the example of someone whose legs begin to ache during a long walk but who ceases to be aware of the aching as a result of his being engaged in a lively conversation. The natural thing to say about this case is that the ache, which is a kind of sensation, continued throughout the conversation even when the person was unaware of it. It remains true that he could have made himself aware of it by suitably directing his introspective attention but there is still a sense in which sensations can fail to be self-intimating. In addition, it is arguable that there are other current mental phenomena 'of which we are not aware, and of which we cannot make ourselves aware merely by the redirection of attention' (Armstrong 1984: 125). One such phenomenon is subliminal perception, 'perception which occurs without the perceiver being aware of it, or being able to make himself aware of it' (Armstrong 1984: 132).

The thesis that thoughts and emotions can fail to be self-intimating is much less controversial than the thesis that sensory phenomena can fail to be self-intimating. Both more and less radical critics of the doctrine of privileged access tend to refer to Freud in this connection. Freud is taken to have shown that the unconscious is a 'really existing thing, exerting causal power' (Armstrong 1984: 131) despite being anything but phosphorescent. If, for example, I can desire something without realizing that I desire it then desire is one mental phenomenon which can fail to be self-intimating. Yet such examples leave it open that the mental is normally self-intimating and that unconscious mental phenomena can in principle be brought to consciousness. This suggests that those who think that the mental is self-intimating have some room for manoeuvre even if Freud's account of the unconscious is accepted.

With regard to alleged infallibility of introspectively based judgements about the contents of one's own mind, radical critics of the doctrine of privileged access deny that any such judgements are absolutely immune to error. Less radical critics allow that there is some introspective infallibility but insist that the scope of such infallibility has been exaggerated in the Cartesian tradition. In defence of the more radical position it might be claimed that 'one may be mistaken about one's own thoughts' (Davidson 1994: 43) and that even introspectively based judgements about one's own sensations can be mistaken. One can think that one is in pain and yet not be in pain. In defence of the less radical position it might be questioned whether mistakes about one's own sensations are really intelligible. In addition, strict cogito judgements appear to be immune to error even if it is not true in general that judgements about one's own propositional attitudes can't be mistaken. As Burge points out, the thought that I am now thinking is both self-referential and self-verifying. In such cases, 'an error based on a gap between one's thoughts and the subject-matter is simply not possible' (1994: 74).

In the light of the infallibility of cogito judgements a blanket rejection of the doctrine of privileged access does not seem warranted. The interesting question is not whether there is such a thing as introspective infallibility but how far such infallibility extends. If only self-verifying judgements are infallible then the fact that few introspectively based judgements about one's own thoughts and sensations are genuinely self-verifying implies that introspective infallibility is not a widespread phenomenon. Yet self-knowledge enjoys other epistemic privileges that are no less interesting. In the first place, one might think that there is an 'overriding presumption that a person knows what he or she believes' (Davidson 1994: 43) and that the possibility that one may be mistaken about one's own thoughts does not defeat this presumption. Secondly, there is the idea that errors about what one thinks or believes cannot be what Burge calls 'brute errors'. Brute

errors do not result from any carelessness, malfunction or irrationality; they do not indicate something wrong with the thinker. In these terms, ordinary perceptual judgements can be brutally mistaken but brute mistakes are impossible when it comes to judgements about one's own thoughts. Finally, judgements about one's own thoughts are direct, in the sense that the knowledge in which they normally issue is not the product of ordinary empirical investigation.

There is much more to be said about each of these epistemic privileges but the important point for present purposes is that they are all privileges that can be enjoyed by judgements that are not strictly infallible. So even if one is sceptical about the idea that self-knowledge is infallible one can think that it is epistemically privileged. How does this bear on Descartes' own position? Although Descartes is often represented as having insisted that self-knowledge is both infallible and exhaustive there is some evidence which points in a different direction. It has been pointed out, for example, that Descartes' thesis that the mind is better known than the body is what Newman calls a 'comparative' rather than a 'superlative' thesis and that Descartes regards introspective judgements about one's own sensations as subject to error. There is also evidence in Descartes' writings of a degree of scepticism about the idea that the mental is necessarily self-intimating. So if a 'Cartesian' conception of self-knowledge is committed to infallibility and self-intimation then it is at least open to question whether Descartes himself was a Cartesian. But we have seen that one can fail to be a 'Cartesian' without going to the opposite extreme of holding that self-knowledge is fundamentally no different from knowledge of the external world. Self-knowledge can be authoritative without being infallible.

If we can know what we are thinking without any empirical investigation how can it nevertheless be true that our thoughts depend for their identities on our relations to the environment? This is a question about the relationship between anti-individualism and the

directness or authority of self-knowledge. So-called ‘incompatibilists’ (Ludlow and Martin 1998) hold that it draws attention to a genuine problem. If I can’t know what the environment is like without any empirical investigation, and my thoughts are individuated non-individualistically, then I can’t know what I am thinking without any empirical investigation. Since I do know what I am thinking without any empirical investigation it follows that anti-individualism is false. In contrast, compatibilists hold that it can be true both that knowledge of one’s own thoughts is direct and authoritative and that some of one’s thoughts depend on relations that one bears to one’s physical and social environment. Even if knowledge of one’s environment must be empirical, it doesn’t follow that knowledge of one’s non-individualistically individuated thoughts must be empirical.

The debate between compatibilism and incompatibilism is still very much alive. If there is anything to the suggestion that Descartes flirted with anti-individualism then it is essential for his purposes that compatibilism is correct. If, on the other hand, incompatibilism is correct, then this strengthens the case for reading Descartes as an individualist. Either way, contemporary discussions of these and many other central issues in the philosophy of mind begin with Descartes. Rightly or wrongly, dualism, individualism, and doctrine of privileged access are all seen as different aspects of Descartes’ philosophy of mind, and the extent to which the philosophy of mind has been shaped by Descartes can be seen in the extent to which responses to these doctrines are seen as responses to Descartes. While few philosophers of mind nowadays would be happy to be described as ‘Cartesian’, and many are explicitly concerned to combat what they see as the errors of ‘Cartesianism’, it is difficult to imagine what the philosophy of mind would look like without Descartes’ contribution.

QUASSIM CASSAM

REFERENCES AND FURTHER READING

- Armstrong, D. M. (1968). A Materialist Theory of the Mind. London: RKP.
- Armstrong, D. M. (1984). 'Consciousness and causality'. In N. Malcolm & D. M. Armstrong. Consciousness and Causality (pp. 103-191). Oxford: Basil Blackwell.
- Boghossian, P. (1998). 'Content and self-knowledge'. In P. Ludlow & N. Martin (Eds). Externalism and Self-Knowledge (pp. 149-73). Stanford: CSLI Publications.
- Burge, T. (1982). 'Other bodies'. In A. Woodfield (Ed.). Thought and Object: Essays on Intentionality (pp. 97-120). Oxford: Clarendon Press.
- Burge, T. (1986). 'Cartesian error and the objectivity of perception'. In P. Pettit & J. McDowell (Eds.). Subject, Thought, and Context (pp. 117-36). Oxford: Clarendon Press.
- Burge, T. (1994). 'Individualism and self-knowledge'. In Q. Cassam (ed.). Self-Knowledge (pp. 65-79). Oxford: Oxford University Press.
- Burge, T. (1998). 'Individualism and the mental'. In P. Ludlow & N. Martin (Eds). Externalism and Self-Knowledge (pp. 21-83). Stanford: CSLI Publications.
- Burge, T. (2003). 'Descartes and anti-individualism: reply to Normore'. In M. Hahn & B. Ramberg (Eds.). Reflections and Replies: Essays on the Philosophy of Tyler Burge (pp. 291- 334). Cambridge, Mass.: The MIT Press.
- Cassam, Q. (Ed.) (1994). Self-Knowledge. Oxford: Oxford University Press.
- Chalmers, D. (1996). The Conscious Mind: In Search of a Fundamental Theory. Oxford: Oxford University Press.
- Davidson, D. (1980). 'Mental events'. In D. Davidson. Mental Events (pp. 207-225). Oxford: Oxford University Press.
- Davidson, D. (1994). 'Knowing One's Own Mind'. In Q. Cassam (ed.). Self-Knowledge (pp. 43-64). Oxford: Oxford University Press.

- Fodor, J. (1981). 'Methodological solipsism considered as a research strategy in cognitive psychology'. In J. Fodor. Representations: Philosophical Essays on the Foundations of Cognitive Science (pp. 225-53). Brighton: Harvester Press.
- Hornsby, J. (1986). 'Physicalist thinking and conceptions of behaviour'. In P. Pettit & J. McDowell (Eds.). Subject, Thought, and Context (pp. 95-115). Oxford: Clarendon Press.
- Kripke, S. (1980). Naming and Necessity. Oxford: Blackwell.
- Ludlow, P & Martin, N. (1998). Externalism and Self-Knowledge. Stanford: CSLI Publications.
- Moran, R. (2001). Authority and Estrangement: An Essay on Self-Knowledge. Princeton: Princeton University Press.
- Newman, L. (2005). 'Descartes' epistemology'. In E. N. Zalta (Ed.). The Stanford Encyclopedia of Philosophy (Summer 2005 Edition). URL: <http://plato.stanford.edu/archives/sum2005/entries/Descartes-epistemology/>
- Parfit, D. (1984). Reasons and Persons. Oxford: Oxford University Press.
- Place, U. T. (1956). 'Is consciousness a brain process?'. British Journal of Psychology, 47, 44-50.
- Putnam, H. (1975). 'The meaning of "Meaning"'. In H. Putnam. Mind, Language and Reality: Philosophical Paper, volume 2 (pp. 215-71). Cambridge: Cambridge University Press).
- Ryle, G. (1949). The Concept of Mind. Harmondsworth: Penguin.
- Smart, J. J. C. (1959). 'Sensations and brain processes'. The Philosophical Review, 68, 141-56.
- Strawson, P. F. (1974). 'Self, Mind and Body'. In P. F. Strawson. Freedom and Resentment and Other Essays (pp. 169-77). London: Methuen.

Strawson, P. F. (1997). 'Entity and identity'. In P. F. Strawson. Entity and Identity and Other Essays (21-51). Oxford: Oxford University Press.