

## *PANORAMAS TEÓRICOS*

CRÍTICA, *Revista Hispanoamericana de Filosofía*  
Vol. XXVIII, No. 82 (abril 1996): 67–107

### WORK DOWN THE MINDS: A SKETCH OF CONTEMPORARY PHILOSOPHY OF MIND

SAMUEL GUTTENPLAN  
Philosophy Department  
Birkbeck College

#### *Introduction*

Should one start by classifying work in the philosophy of mind in terms of phenomena or problems? The choice determines whether one comes to see the subject as primarily a set of problems illustrated by particular features of the mental realm, or as a set of phenomena in that realm which give rise to domain-specific problems. In the end it may not matter much where you begin, but I think there are definite advantages to starting with the phenomena. In particular, one can better understand the problems, and how they fit together, if some attempt is made at the beginning to survey, in as neutral a way as possible, the things which count as specifically mental.

If you ask those best placed to know such things, i.e., possessors of minds, it would be apparent that the phenomena divide into three broad categories: attitudes, experiences and actions. Taking the major part of my task as defined by these categories, I shall in each case try first

to say something uncontentious about them, and about the more specific phenomena they include. The aim here will be one of orientation. Having done this, I shall then describe how philosophy has got to work on each of them.

Two notes before I begin. First, the philosophy of mind is currently the most active area in all of philosophical research, and it has connections to psychology, computer science and those related disciplines which go to make up what is now known as ‘cognitive science’. Sheer volume thus makes the task of summarizing the philosophy of mind, and its current obsessions, not merely difficult, but, as I have realised, virtually impossible. I have had to be selective in matters of detail, though I have tried to touch on most of the central issues. Still, the only way to get properly to grips with the subject is through further reading in the literature.

Second, rather than cluttering up the text with citations, I have chosen to save most of the bibliographical references for the end. There you will find lists of reading which, whilst far from compendious, includes many works which themselves are heavily laden with bibliography.

### *Attitudes*

The attitudes count amongst their number believing, desiring, intending, fearing, regretting, hoping. . . The list is certainly not numberless but it is very large. Russell called these items ‘propositional attitudes’ because he thought it obvious that they were attitudes towards propositions. It turns out that this is far from obvious, though the name has stuck. In any case, it seems harmless enough given the fact that the canonical description of a propositional attitude —what is usually counted as the standard way of ascribing or attributing such an attitude— is:

$X$  believes (desires, intends, hopes, fears, regrets) that  
 $p$ ,

where ‘ $X$ ’ is some agent and ‘ $p$ ’ some indicative sentence. Given that indicative sentences are pretty good ways to express propositions, we get at least an initial justification for Russell’s name. But, of course, since ‘ $p$ ’ is an indicative sentence, there are those who wonder whether ‘sentential attitude’ might not be more revealing. They will get their say shortly. First, though, I should like to point out that, if we think about the attitudes without allowing ourselves to be swept along by Russell’s name, it would be obvious that there are cases in which attitudes are reported (attributed, ascribed) without the benefit of sentences, whether or not expressive of propositions. Thus, one can ‘desire  $t$ ’ where ‘ $t$ ’ is the name of an object (or person) and one can ‘intend to  $a$ ’ where  $a$  is some act. It is an underdiscussed question whether object and infinitival constructions are mere stylistic variants on the more basic sentential form, or whether they mark some interesting difference in attitude. Evidence that something interesting might be going comes from the fact that certain attitudes do, and certain others do not, allow themselves to be attributed in the different styles. For example, I can desire  $t$ , to  $a$  and that  $p$ , but, unless I change the subject, I cannot believe  $x$  or to  $a$ .

An absolutely crucial feature of the attitudes is that the indicative sentences used in specifying them need not be true even though the whole attribution may itself be true. Thus, one can truly believe that London is at the same latitude as New York, though London is certainly not at that latitude. There are exceptions to this: one cannot know, guess, remember that  $p$ , unless  $p$  is true. But by and large the attitudes are not, as is said, ‘factive’, and this has enormous consequences for our dealing with them. Under its most traditional name, this feature of the attitudes gives

rise to the ‘problem of intentionality’. It is in part the problem of explaining how we can *say* that someone could be truly related to what may well be non-actual, and it is also the problem of explaining how someone can *be* so related. (As will be discussed below, there is a tendency to shift between the problems generated by the words we use to ascribe attitudes, and the nature of the ‘states’ so ascribed. Both sets of problems are very difficult.)

Attitudes are ascribed to others largely on the basis of what they do and say, and to ourselves on what can seem no evidential base at all. This does not entail, as some have thought, that we cannot be wrong about what we believe, desire or intend; far from it. Even without the benefit of psychoanalytical theory, it seems perfectly possible for someone sincerely to claim, for example, that he intends to help his friend, when someone else —perhaps the friend— might be even better-placed to doubt that there is any such intention.

The partial list of attitudes given above is shapeless, and there is inevitably a tendency to organise them in some way. Typically, believing (and knowing, though the latter seems to many to be an attitude ‘plus’ something non-attitudinal), thinking, inferring, confirming, expecting, amongst many others, are seen as cognitive attitudes. The rest are often treated as merely non-cognitive, but it is possible to separate them into broadly two further categories. Desiring, wanting, needing, and the like are the conative attitudes; and fearing, loving, hating, regretting, envying, etc. are simply called ‘emotions’. Classification here is not a precise business. Without endorsing it, there is a view which has it that the cognitive attitudes are those which reflect in some way *how the world is*, whilst the conative attitudes express a picture of *how the world should be*. It is often said that the former have a ‘mind-

to-world' direction of fit, whilst the reverse is true of the latter.

Given this sort of classification of attitudes, emotions tend to be embarrassments: on the one hand, they often presuppose lots of things about the way the world is: to be proud that  $p$ , I must have had something to do with  $p$ 's being the case. On the other hand, they seem somewhat appetitive: my hating that  $p$  is going to be bound up with my doing what I can to make not- $p$  the case. Much more attention has been paid to the emotions in recent philosophy, but it is not always considered part of mainstream philosophy of mind.

Differences among them aside, attitudes are stranger than they might at first seem. When one says, for example, that Smith believes that global warming has begun, one is certainly trying to say something true about Smith, and undoubtedly if Smith was not the sophisticated, mind-possessing object he is, we would not be in position to make such a claim. Part of what it is to have a mind is to be a fit subject of attitude attribution. But it is both very unclear what exactly is being said about Smith in such a claim, and what is the point of making it. I assume we know (approximately) what it means to say that global warming has begun. But how does this knowledge fit in to the strange practice of using it to characterise Smith's mental state? (Being careful about this use of the expression 'mental state' will be important, but not yet.) And, just as puzzlingly, why do we bother to say such things?

It is crucial to recognise that the questions we are here concerned with do not yet involve the so-called mind-body problem; the wonder is not how it could happen that creatures who are essentially complex physical systems could have beliefs, desires and the rest. Rather, they are prior questions about the form and point of the words used in attitude ascriptions. To be sure, answers to these questions

are bound to have consequences for our view about how attitudes fit into the physical scheme of things; and, if we have some inkling about how attitudes do or could fit in to the physical world, this will almost certainly constrain our answers to the original questions. But the questions themselves make sense, and are pressing, even if we did not care much about the physical credentials of attitude attributions.

The first of these questions, which is essentially about the form of words used in attitude attribution, has had a very long history, much of it outside what is now considered the philosophy of mind. For ages now, philosophers of language and logicians have been working on ways to accommodate attitude constructions, and their efforts have been constrained by the needs of semantic theory rather than by any interest in the place of attitudes in the physical world. The problem, in outline, is as follows. Considering attitude attributions to be true claims about individual human beings, they must (apparently) respect certain semantical rules of substitution and inference. Unfortunately, they do not seem to. If it is true that Lois Lane believes that Clark Kent works in her office and, given the true identity of Clark Kent with Superman, it should also be true that Lois Lane believes that Superman works in her office. Yet this is something that she failed to believe in spite of what most followers of the Superman stories regarded as massive evidence. Similar problems can be made to arise in respect of logical inferences involving attitudes and other sorts of substitution.

The ingenuity which has been devoted to these difficulties in logic and the philosophy of language is impressive. Frege has some claim to being the first to highlight the problems, and Quine kept the ball rolling. (He, unlike Frege, had hoped that, in some way or other, people would take the semantic problems of the attitudes as rea-

son enough to keep clear of them; one could say that he wanted the ball to roll over the edge.) Amongst the kinds of solution that have been proposed are: detecting ambiguity in attitude ascriptions (so-called relational and notional readings); finding a fundamental bifurcation in attitudes between those which relate to sentences or propositions (*dicta*) and those which relate to objects themselves (*res*); finding attitudes to involve relata that are hidden from view in the apparently binary relation, ‘*x* believes that *p*’; finding alternative ways of describing the propositions that towards which we take up attitudes. And this is only a partial list.

It is of course not really possible to solve the semantic problems of the attitudes without at some time asking oneself exactly what an attitude is. On one common view, a smooth-functioning semantic theory should tell us how the semantic features of the parts of an attitude sentence contribute to the truth conditions of the whole. But, whilst you can work on this project without worrying overmuch about the truth condition of the whole —intuitions about which attitude sentences are true/false seems enough to get started— at some point deeper questions about such truth conditions need to be addressed. In particular, we must ask ourselves the more metaphysical question of what truth, if any, about the world is expressed in a typical attitude sentence —what must the world be like if ‘Smith believes that global warming has begun’ is true of it? (In any case, it has become apparent that our intuitions about attitude sentences are actually quite malleable. It is no longer easy to get philosophers to agree to its being false that Lois Lane believes Superman works in her office, or, more importantly perhaps, to agree that this is what ordinary intuition tells us. And if you read widely enough in the literature on the semantics of the attitudes, you will find yourself losing your grip on what you may have thought before you were

presented with the many intricate variations on this sort of example.)

The metaphysics of the attitudes is of course inseparably bound up with the question of how a physical system, namely a human being, can be a believer. This marks the appearance of the traditional mind-body problem as it impinges on the attitudes, and I shall say more about it below. But first I should like to consider the second of the questions I raised earlier, viz., why do we bother to make attitude attributions to ourselves and others? You might have thought this question would have figured historically ahead of the semantic and metaphysical problems, but surprisingly (to me) it did not. In Quine 1960, there is brief (and characteristically dismissive) speculation about what we are doing in using these attributions, but the full-scale debate has only been more recently joined. This has always surprised me because, given the uses to which attitude attributions are put in history, social theory, economics, as well as of course in everyday chat, it is not unreasonable to regard the study of propositional attitudes as foundational for the whole of social studies.

What no one doubts is that the practice of ascribing attitudes is a universal feature of human behaviour. But that said, there is very little agreement about the practice itself. If one assumes that, in making attitude ascriptions, we aim to say true things about each other, this leaves us still short of an understanding of why these truths matter so much to us, why they figure so fundamentally in human behaviour, why an individual's inability to engage properly with this practice constitutes a deep disturbance of some sort. The whole package of attitude attribution has been called 'folk psychology' and the name, however flawed, has stuck. I believe that Dennett was the first to coin the term, though he has told me that he thinks he got it from someone else. In any case, he made it part of the



contemporary philosophical lexicon. It is flawed because of what it unsupportedly intimates about the practice of attitude attribution. In particular, it suggests that the practice is both a kind of psychological theory and a naive one at that; that it is a primitive attempt to deal with human behaviour matching in ineptitude attempts to deal with certain other phenomena such as folk medicine and folk physics. As shall be seen, however, there is no shortage of those who resist either or both suggestions.

Beginning with the suggestion that folk psychology is in some sense a theory, we can draw the following picture. Human behaviour is describable—in terms drawn from folk psychology—as intentional, as consisting in actions. (The problem of action will figure in the third part of this article.) These actions are made intelligible to us (or understood, or explained, take your pick for now) by our being able truly to ascribe various attitudes to the relevant agents. Thus, my buying a newspaper this morning is explained, etc., by my wanting to know who won yesterday's election in Australia and by my believing the newspaper would give me that information. (A whole range of my background beliefs, desires and other attitudes are left out of this picture, but the example is typical of those philosophers use.)

Everything of importance here turns on the notions of understanding, intelligibility and explanation. At one extreme is the view that folk psychology is indistinguishable from the kind of explanatory theory one finds in scientific contexts, except of course in its sophistication. Concepts like belief, desire and intention are likened to the concepts of a scientific theory such as mass and energy. By deploying these concepts in making true statements about some individual, we construct a mini-theory of that individual's behaviour, and, if we get it right, we have thereby explained the behaviour. In this way, our interest in folk psychology is not difficult to justify: bound up as we are

in the behaviour of those around us, and no less concerned to make sense of our own activities, we naturally appeal to what seems the most plausible way to describe and explain what would otherwise be a confusing mass of 'data'. Moreover, insofar as the key notion here is that of explanation, and since, in scientific contexts, such explanation is bound up with finding causes and/or causal regularities, this picture of folk psychology invites us to think of the mental states we attribute as themselves causes of behaviour.

Less strict views of the suggestion that folk psychology is a theory are possible. Perhaps in the scientific sense, a theory is some body of claims employing a set of theoretical concepts which promises to yield explanations, in some way causal, of a range of phenomena. But in a looser sense, we can think of a theory as simply a way of understanding or making intelligible the phenomena. In describing behaviour as intentional perhaps the point is not so much theoretical explanation as it is a kind of locating of the behaviour in a context where it can be seen as rational, as done for reasons. Such placement makes irreducible appeal to our ability to understand, from the inside, so to speak, what a reason is, and to have some idea of what makes behaviour rational. In this way, it does not match a scientific account —an account whose explanatory aims encompass the behaviour of systems from the outside.

The idea that intelligibility or understanding is the aim of folk psychology, rather than quasi-scientific explanation, can be seen as at the same time a reason for moving away from the whole idea of folk psychology as having any kind of theoretical pretension. Further reasons come from some awkward questions that the so-called 'theory theory' of folk psychology has had to answer. On the one hand, if what we are doing in attributing attitudes is something like what goes on in science, how does it come to be part of

the repertoire of normal adults and not just of those particularly curious ones who like to explain things? And, on the other, what about all the evidence that has accumulated showing a pattern of acquisition of folk psychological notions at the age of about 4 or 5?

In response, there has grown up in the past ten years a completely different paradigm for understanding what goes on in folk psychology. Called the ‘simulation theory’, the core idea here is that, when a normal human adult sets about attributing attitudes to another, what he or she does is not so much frame a theory of the other—a set of claims—but rather makes essential use of his or her own mental capacities to simulate the mental state of another. The idea could be traced back to some of the things Quine (1960) says about attitude ascription. There he describes the practice as depending upon our ability to cast our real selves into unreal roles: saying what Henry believes consists in imagining yourself in Henry’s place and then saying what you believe.

In the very recent literature, there has been a flourishing discussion of just what this simulation theory is committed to, and whether it has importantly different consequences for our understanding of the mind from the theory theory. On the one hand, there are those who regard the simulation theory as merely a suggestion of how we manage the knowledge contained in its more explicitly theoretical cousin. And, on the other, there are those who claim it to be quite disjoint from the theory theory. Some even suggest that the simulation theory has important connections to debates about the nature of historical explanation that took place earlier this century: the simulation theory is seen as a close relative of the ‘*verstehen*’ account of social and historical theory.

Whatever one ends up saying about the point of folk psychology, the metaphysical issues described earlier are

just beneath the surface. If you think that theoretical description is the key, then some account must be given about how attitude claims connect to the world so we can plausibly regard them as pulling their weight. Think here about the typical scientific theory, say the theory of the chemical bond. Claims made in it about the number and position of electrons in atomic orbit determine the chemical behaviour of various substances. The theory not only works, it is generally regarded as (something approaching) true, because we are pretty sure about the relationship between claims in chemical theory, and the behaviour they explain, and more basic claims in physics about the atomic nature of matter. Simple-minded though it may seem, it is commonly said that the theory of the chemical bond is right because it describes the real properties and relations that obtain amongst the substances which figure in chemical interactions. If on the other hand, you think that theoretical explanation is not the point of folk psychology, or if you favour some version of the simulation theory, you still cannot avoid the metaphysical question. Whether it is a matter of understanding how there can be reasons in nature, or what it is about organisms like us that makes us suitable models for attitude simulation, one is ultimately asking how the practice of attitude ascription can be grounded in the world of physical objects and properties.

A note of caution should be inserted here before I discuss some of the directions that have been taken in answering the metaphysical question. The mind-body question, as it traditionally figured in metaphysics, was relatively light on any detailed characterisation of the mind. The general assumptions seemed to be that: human beings (at least) have minds; that minds have properties which are *prima facie* difficult to reconcile with the naturalistic or, somewhat more restrictedly, physicalistic conception of what goes on; and that an unacceptable dualism lay in wait for us if we

could not nonetheless reconcile our mindedness with our physicality. In contemporary philosophy of mind, there is a much greater awareness of the mental landscape, so each category of the mind faces the demands of physicalism in slightly different ways. Hence, though some of what follows will recur in my later discussion of experience, I do not think this repetition is necessarily a bad thing. The problems for physicalism raised by the attitudes are sufficiently different from those raised by other phenomena of mind to make it more revealing to discuss each set in its proper context.

The attribution to, say Smith, of an attitude is often thought of as a claim to this effect: Smith has a mental state (believing, desiring, intending, etc.) with a particular content (that global warming has begun, that today is Tuesday, etc.). Taking ‘mental state’ to be no more than a general term for the specific attitudes, and taking ‘content’ to be simply a way of referring to whatever is believed, desired, etc., this is a fairly non-committal way putting it. However, the expression ‘Smith’s mental state’ conjures up a picture of Smith as having or possessing some intrinsic feature. Think here of claims about Smith’s state of health: to say that Smith has a belief could suggest something like the claim that he has a cold. (I do not, by the way say this way of looking at it is wrong. Just that we must be careful not to insist on it because of the ease with which we slip into this way of talking.)

If beliefs are like colds, then we have a pretty shrewd idea where to look for them in the physical world: they will figure, at some level of description, in the physical make-up and context of the individual believer. But of course in one important way beliefs are not like colds: they have contents. This seems obvious enough from the form of words we use to attribute them. Moreover, difficult as it has been for philosophers of language to regiment attitude ascriptions,

the problem is intensified when we try to understand how the attitudes themselves —the ‘states’, which it is natural to say, realise the attribution— fit into the natural order. How can there be states of nature which involve a relationship between an individual human being and some possibly non-actual state of affairs that is cited in the content-specifier of an ascription? And how can the contents of attitudes play a role in the push-pull flow of the causality? Saving the title ‘problems of intentionality’ as a general name for the tangled web of issues raised by the attitudes, I prefer to call this the ‘problem of naturalizing intentionality’.

There are many different ways to classify the suggestions that have been made for this project of naturalization, and most do not do justice to the nuances of views actually held. Still, it helps to have some broad scheme of classification, and my preferred scheme begins with the following question:

What *kind of thing* is a belief?

There are those who think we can best begin to answer this question by looking around for the physical state which corresponds or, more committedly, is type-identical to the belief. For them, the question is very much like the question: what is a bicycle?, in that it invites some story about the physical nature of the relevant item. But, however obvious this strategy may seem, there is another one. Sometimes when asked the ‘what is...?’-question we do not immediately reach for physical descriptions. For example, asked what a word processor is, we would most likely answer by citing the characteristic job such devices are meant to perform. Call this indirect approach to the above question a *functional*, as opposed to a physical, specification. (As will be discussed shortly, the functional view is only one of a range of what I have called ‘indirect’ answers to the above question.)

Indirection in answering the earlier question does not require denying that a word processor is a physical device. Rather, given that there are so many different physical objects which do word processing, it would be more informative at least to begin with the functional specification. In the case at hand, it is stretching the point only a little to speak of a functional specification of a belief. Think of the typical sources that a belief has in an individual subject, and the typical consequences there are of having that belief. The functional role of a belief, then, is that item in the mental economy of a subject which comes into being as a result of certain complex perceptual and mental sources ('inputs') and leads to an equally complex pattern of behaviours, dispositions to behave and perhaps other mental states (all these are 'outputs'). In the end, the 'what is...?'-question is answered by citing (or advertng to) the imagined functional role of the state in question.

A special appeal of this idea of functional specification is that it fits comfortably with the now widespread idea that the human mind either is, or is analogous to, a computational device. Ask the 'what is...?'-question of a specific piece of computer hardware and you will get some physicalist answer very much along the lines of the 'icycle approach' to the question. But ask the same question of some part of a computer's program and the answer will be functionalist: 'an index cycle takes a process as input and outputs an increase of one into a counter whilst restarting the process.' That is what an index cycle is, and it is laid out without reference to any piece of hardware.

Computational models are now at the centre of much work in both philosophy of mind and psychology. Indeed, in the latter field, there is a tendency to take the existence of some computational model as a background assumption of research. (A current hotly debated issue in philosophy and cognitive science concerns the nature of the compu-

tational model which seems most plausible, whether it is more like the traditional ‘symbolically’ programmed device, or is instead connectionist.) However, there are interestingly different ways of answering the ‘what is...?’-question indirectly, and not all of them encourage computational modelling.

Suppose, for example, that we thought of the web of attitudes, not as systematic in the way required for computational-functional talk, but as organised according to the demands of rationality and intelligibility. This would require our appealing to the notions of rationality and intelligibility as a way of understanding what an attitude is, and it therefore counts as a way of approaching the original ‘what is...?’-question in other than immediately physicalist terms. Instead of computational or functionalist, this approach could be called *attributionist* or *interpretationalist*. Roughly, on this view, the metaphysical question of what an attitude is can only be answered by appeal to the practice of interpreting others as having beliefs, desires and the rest. Smith’s belief that  $p$  is whatever the best theory of interpretation —the one answering to the demands of rationality and intelligibility— ascribes to Smith. Does this mean that Smith’s belief is not a physical thing? The interpretationalist here smiles inscrutably: ‘I never said that an attitude isn’t a physical thing. In fact, given that everything is almost certainly physical, so are attitudes. But you don’t *understand* them by describing anything about the physicality of those to whom they are attributed.’

How can an attitude be a physical thing, even though we don’t understand it as physical? In one way or another, the notion that is appealed to here is *supervenience*. Smith’s belief that  $p$  is said to supervene on physical facts in or around Smith, and, if we get this notion right, the hope is that we will be able, as physicalists, to sleep soundly



without having to go as far as thinking that Smith's belief is intelligible as a physical item.

Asking 'what kind of a thing is a belief?' should only be seen as an icebreaker. A really difficult follow-up question is:

And what does its content do?

As described above, indirect accounts suggest that the first, or perhaps central, way to understand what an attitude is is not as a physical thing. But, since there are few self-proclaimed dualists around, virtually everyone — computationalists, functionalists, interpretationalists— insists that attitudes are physical in some sense. And this makes the above question especially pressing. Suppose for example, that you think of beliefs and desires as, at some level of description, causes as well as reasons for actions. Then, given that beliefs and desires have contents, we would expect some account of how contents play a role in the causal flow from mental state to action. But how can contents play such a role? After all, if my belief that  $p$  supervenes on some physical feature of the world —or if it is token identical to some physical state— then one presumes that that state does the causal work. As there seems to be no place for 'content' in purely physical states, and as causality requires the robustness of the physical, content drops out as 'inert' or 'epiphenomenal'. In recent literature, this issue has been widely discussed.

One suggestion about the attitudes which promises to allow a proper place for content, as well as supporting a broadly computationalist account of the attitudes, has attracted a large following, especially amongst psychologists. Central to this suggestion is the idea of what is called a 'Language of Thought'. Sentences in a natural language can be understood in two ways: as inscriptions, i.e., physical objects of some sort; or as expressive of meanings, as having

contents. Trading on this, the Language of Thought (LOT) hypothesis suggests that, corresponding to contentful attitudes, there are tokens (inscriptions) of these sentence-like attitudes written in our ‘mind/brain’. (Exactly at what level of description we are to conceive of LOT inscriptions is an open question, but they are certainly to be regarded as physicalistically respectable.) Putting these suggestions together with a broadly functionalist account of the differences between one attitude and another, the picture is this: a belief that  $p$  is a type of functional state of the mind/brain in which an inscription of a LOT sentence meaning that  $p$  figures as an essential element. There are various possible explanations of what it is for some such sentence to mean that  $p$ , one of which is broadly causal. A LOT sentence means that  $p$  when it is appropriately and reliably connected to an appropriate state of affairs.

The LOT thesis does not necessarily require a causal account of content, and there can be accounts of content which do not go so far as to insist that there is a LOT. In the recent literature, there has been a growing debate about whether we can give some ‘historical’ account of content, perhaps even an evolutionary one. Thus, for a state of some individual to be a belief that  $p$  it might be necessary for there to be in place some historical link—in the relevant individual— between that state and  $p$ . Or perhaps the link is not intra-individual, but is a feature of the evolution of the species of which the individual is a member. The range of positions here is enormously complex and no summary can provide the subtlety required.

### *Experience*

What I call the category of experience is probably central to most peoples’ conception of the mind. Indeed, there is a tendency to think of this category as virtually defining

mindedness. Typically included here are the experience of pain and other bodily sensations, awareness of our trains of thoughts, intentions to act, emotions and moods, and, perhaps most difficult to characterise accurately, our sense of what it is to perceive colours, sounds, tastes, etc.

It is somehow natural to think of these phenomena as specially presented to a subject in various ways. Thus, to begin, as most discussions do, with the example of pain, it is not difficult to convince the newcomer to the philosophy of mind that pains are self-intimating, that our experience of them is incorrigible and that they are ultimately private to their sufferers. More concretely, the thought is that if I am in pain then: (i) I cannot fail to notice it (this is self-intimation); (ii) if I sincerely say that I am in pain, then I could not be wrong (knowledge of my own pain is incorrigible); and (iii) I cannot in any real sense share my pain with anyone else (it figures in a private mental landscape).

Pain is more prominent as an example of experience than it perhaps ought to be. There are reasons for thinking that it is not wholly typical of the category of experience, and, by treating it as such, we are storing up more trouble for philosophy than is strictly necessary. Hence, it is important to look closely at other things that come into this category.

Consider first the experiential aspects of perception. When we look directly at a yellow wall with our eyes open and in a good light, we will usually be able to judge that the wall in front of us is yellow. Perception in appropriately satisfactory conditions seems a good source of knowledge about our surroundings. Now most people can be brought to agree that somewhere in the process of our making such perceptual judgments is a place for a mental occurrence which is not itself *of* the wall or its colour. Traditionally, there have been a number of (possibly non-equivalent) ways of trying to draw attention to this occurrence. It has been

said that we have, or have presented to us, a sense-datum as of the yellow wall; that we have a visual sensation of the yellow wall; or that our perception of the yellow wall has inner qualities in terms of which we come to make our judgment. Language is a problem here. The predicate 'yellow' is certainly not something we think appropriate in a description of any inner sensory occurrence. Even though colours have sometimes been metaphysically downgraded to the status of secondary, mind-dependent, qualities, there can be little doubt that what passes through my mind when I see a yellow wall is not itself something yellow. Yet, there really seems to be something going on, as it were, inside, and given that it is, as we think, available to consciousness, there ought to be something we can say about it.

Additionally, there is a wide range of experiences which are not implicated in our sense-perception of the world or our awareness of our bodily states. When I get some bad news—a paper of mine which I had thought rather good is rejected—I may be depressed, or 'down'. If I have been under a great deal of pressure at work, I may feel annoyance with a colleague who makes yet another demand on my time. More happily, it can sometimes happen that my thinking results in, as one says, an idea. Finally, when I set out to do something, my action seems something of which I am aware and which I can keep on track by paying attention to what I am doing. Each of these is an experience, something which passes through, or resides in, my consciousness, and there certainly seems to be 'something it is like' to have them.

Sensations of pain, or perceptions of the 'outer' world, seem to share a feature with the moods and thought processes just described: we have little or no vocabulary with which to characterise them directly. I can, as I have done above, say what brought them on, or what they lead to, but I seem unable to convey their felt quality to anyone else

other than by metaphor. Moreover, even if we are resistant to the idea that items of experience are wholly ineffable, there does seem to be a kind of asymmetry in our knowledge of them: my knowledge of my own experiences and conscious states is authoritative in a way that contrasts with my knowledge of other things.

Items in the category of experience are, to put it mildly, very difficult for philosophical and psychological theorizing, and it is not uncommon for writers who have a lot to say about the attitudes, to back off when the subject is consciousness or experience. Difficult as it is to see how contentful states, the attitudes, can be accommodated within a naturalistic or physicalistic world-view, there is no shortage of suggestions as to how the trick might be turned. But consciousness brings us up short. There seems to be what is generally known as an *explanatory gap* between what we think the physical world is like and what we think of experience. For this reason, argument in this area can sometimes seem more concerned with the gap than with accounts of the rich range of experiential phenomena themselves, and this has tended to make the debate seem strangely indirect. (Though, as will be discussed below, there has recently been a greater concentration on the nature of consciousness itself.)

There have been four responses to the explanatory gap: (i) some cannot see that there is such a gap—for them, experiential phenomena simply are states of the brain; (ii) some find in it a reason to doubt the reality of the phenomena of experience; (iii) some think it a reason to give up on the physicalist world-view; and (iv) some think it shows just how limited we are as cognitive beings.

In considering these responses, one must begin with the doctrine of physicalism itself. If nothing else, the stresses put on this doctrine (or doctrines) by issues in the philos-

ophy of mind have served a useful purpose, allowing us better to understand the nature of our commitment to it.

At its most strident, physicalism consists in the claim that the objects and properties which make up the world are those uncovered by a mature physics. Anything counts as real just to the extent that it can be shown to be reducible to these fundamental physical objects and properties. If there are chemical properties, then they are in principle reducible to physical ones; if there are biological categories, then they too are physical at some level of description. And if the phenomena of the mind are to be counted as real, then in some way or other they must be seen as congeries of physical properties and relations. Many who count themselves as physicalists do not accept this version of the doctrine, but it is crucial to understanding weaker, or less strident forms, that we begin with the most robust.

Less committed physicalisms come from focussing on the idea of reduction. Perhaps the world is physical, in the sense of being composed solely of physical things, but there seems room to believe this whilst resisting the idea that we can understand, explain, make intelligible everything that happens in physical terms. Cleaving to the idea that the world is at bottom physical, but combining this with a pluralism of explanatory objects and properties—a plurality of what counts as real—promises to allow a kind of irreducible reality to biological and mental phenomena. Of course, as was seen in the case of the attitudes, there is a lot of work to be done before we have a right to this position; attitudes have contents and it is far from clear how these can be reconciled with the demands of even a pluralist realism.

As noted above, the obstacles to counting the phenomena of consciousness among the real properties of the world are considerably greater than they were in the case of the attitudes. For even if we are satisfied with a physicalism

that does not insist on reduction, it seems impossible to grasp how consciousness could so much as be composed of something physical. For those not clear why this should be so, there are two much-discussed arguments intended to back up the initial intuition.

(i) One could imagine beliefs, desires and other contentful states being handled by some computational-functional model, but still be very short of what is needed to understand experience because it is perfectly conceivable that two beings could be functional replicas of one another even though one had experiences and the other lacked them. This is because, as we saw, functional states are understood in terms of their characteristic input/output relations, but experiences have qualitative properties —‘qualia’, as they are now universally called— which do not themselves figure in the nexus of functional inputs and outputs. This line is called the ‘absent qualia’ argument. Somewhat to the side of this, perhaps a corollary, is the ‘inverted qualia’ argument, which many regard as somehow reinforcing the first. Here one is asked to imagine two functionally identical human beings whose qualia are inverted in respect of one another. Being functionally identical, it would be impossible to separate them on the basis of their discriminatory interactions with the world and each other, but when one perceives a red object, she experiences the very qualia that the other has when perceiving something green. So long as this is coherent, and this is of course a fiercely debated issue, it would seem that the experiential phenomena —what it is like to see something as red or green— float free from the causal network that underlies at least a large part of what is regarded as mental.

(ii) It has been said that the absent (and inverted) qualia arguments are not against physicalism *per se*, so much as against one particular account of the mind, viz., function-

alism. In particular, it is claimed that, if consciousness is treated as identical to, or composed of, physical states of the brain, then these arguments lapse. However, this line seems to me to mistake the background assumptions. Given that we just cannot grasp how our experience of seeing red or of having a pain could be a brain state, some functionalist (or other indirect) account of the mind offers at least some hope. The aim of the absent and inverted qualia arguments is to dash that hope. Still, the second argument—the so-called knowledge argument—seems a more direct attack on physicalism.

This argument begins with a thought experiment. A scientist, Mary, is unable to see colours. But she lives at a time when our knowledge of colour perception is so advanced that she is able to know everything about the way the brain processes reflected light, and produces colour responses in human subjects. One could even imagine that she has invented a prosthetic device which uses this information to predict with complete accuracy the colours of objects; with the device she can say what colour anything is, and do so in a way which precisely matches the judgments of any normally sighted human being. Now one day she acquires the capacity to judge colours without this device. Before this happened, she knew that the wall was yellow; but afterwards, the argument continues, she knew in addition a further fact, namely what it was like to see yellow. The conclusion we are encouraged to draw is that, since certain facts are not caught in what is, by hypothesis, a complete account of all the physical facts, physicalism—intended as it is as the whole story about the world—is wrong.

The above arguments depend, in effect, on the conception of the mind which began this section, some parts of which I, perhaps rashly, counted as commonsensical. However, in many different ways, it has been argued that the picture of the mind as an inner landscape of which we can



have incorrigible and authoritative knowledge is far from mandatory. Indeed, it has been suggested that, were it not for the writings of Descartes, we would not even regard it as such. In broad outline, the attack on the so-called ‘cartesian’ picture of the mind comes from two sources: one is inspired by the work of the later Wittgenstein on the ‘private language argument’, and the other arises more directly from the explanatory gap described above. It is important to keep these sources distinct, even there is inevitably some sharing of argument.

Whilst it is always difficult to say anything about Wittgenstein without being accused of some misunderstanding, I will venture this much: he did not base his attack on the cartesian conception of the mind on any prior commitment to physicalism. Indeed, he probably had no more time for the physicalist world-view than he did for the cartesian conception itself.

Wittgenstein worried away at the idea that we could understand the inner ‘world’ of experience without seeing it as embedded in a world of human subjects acting in inter-subjectively available environments. The elements of this attack are, as noted above, highly controversial and there is no substitute for studying his texts and their interpretations in detail. However, it does seem clear enough that his position should not be understood as a kind of behaviourism. The phenomena of experience, and the mind generally, are not reducible to the complexes of action we appeal to in attributing them. (There are even hints of a kind of epiphenomenalist dualism in Wittgenstein, but I would get into trouble if I pursued this too far.)

The second attack on cartesianism comes from those who, seeing no way to bridge the explanatory gap, find it more congenial to cast doubt on the reality of consciousness than to doubt the physicalist world-view. Sometimes this attack takes the form of a direct assault on our highly

metaphorical description of the inner landscape with private objects and properties. Here there is a clear overlap with some of Wittgenstein's arguments. But, in the main, the source of the arguments has tended to be recent psychological research suggesting that we are actually a good deal less in touch with our conscious thoughts than we might have naively thought. I suppose the logic is this: if we are not as well-placed to assert what is going on 'within' as we had thought, then perhaps the cartesian picture has deeper flaws than its metaphoricality. For example, whilst we tend to think of pain as a bodily sensation incorrigibly and privately presented to the sufferer, which is the cause of certain of our behaviour, it has been suggested that elements of this picture are less coherent than we ordinarily suppose. The claim is that there could not be an 'inner' state with the right connections to external world. Similar arguments have been mounted against the qualia supposedly found in our perceptual experiences. And, most recently, there has been a good deal of discussion about whether our conscious intentions to act play much of a role in our actually doing anything.

This sceptical assault on the cartesian conception goes hand in hand with what has been a very recent upsurge in the study of consciousness, an upsurge which has taken place in spite of the acknowledged difficulties of dealing with this elusive group of phenomena. Psychologists have tended to be cautious: phenomena of attention and the planning of action have figured more centrally in their work than such items as the experience of pain. But there has been an interesting coming together (almost) of psychologists and philosophers over the study certain aspects of perception. The initial focus for this research has been the phenomenon of 'blindsight'. Apparently, and this word must be taken seriously, it seems possible for perception to take place in the absence of any conscious awareness. Sub-

jects who are blind and have no capacity for the kind of inner visual sensation described earlier, seem nonetheless able to discriminate visually between various presentations. It has also been claimed that perception without experience can occur in normal subjects under the right experimental conditions.

Exactly what consequences this has for philosophical accounts of perception and perceptual experience is currently controversial, and it is still too early to say what, if any, effects this debate will have on our overall conception of experience. However, one noticeable outcome of these debates has been their acting as a stimulus for philosophers to make distinctions, a sure sign of philosophical life. Perhaps the most central one is that between *phenomenal* and *access* consciousness. Roughly, phenomenal consciousness points to that elusive qualitative aspect of experience, that sense of what it is like to have a pain, an intention, or to perceive the world from a point of view. This is admitted by all who think there is such a phenomenon to be the most difficult to understand. In contrast, access consciousness is that —possibly non-phenomenal— presence that our beliefs, desires and other contentful states have. Moreover, there is currently some hope that we are making progress with it. It is said that one has access consciousness to some contentful state when one is able to report on it. For example, I have no doubt that there are features of my perceptual system, as I look at the screen in front of me, of which I am completely unaware. Somehow I am able to process the incoming visual and linguistic material, but in ways which I could not tell anyone about because, as one says, I am not conscious of them. However, I am conscious of, for example, the length of the paragraph I am currently writing; my belief that it is getting longer than I would like, is one to which I have access, and which can therefore enter into my plans.

There are further distinctions to be made in respect of this general notion of access consciousness, but, even in rough form, it is possible to see how it has been used to tame the more difficult case of phenomenal consciousness. In broad outline, the idea is that phenomenal consciousness is nothing other than our having access to experiences which otherwise seem so elusive. My experience of a pain consists in my having a requisite thought about it; possessing a point of view and a sense of what it is like to have various experiences, consists in a kind of monitoring of what is going on, where such monitoring is itself understood in terms of something like access consciousness.

It is easy enough to see how the attempt to understand experience in terms of some notion of access consciousness can dredge up the cartesian picture and the explanatory gap it seems to induce. Some insist that, in trying to define consciousness in terms of access consciousness, one has actually left out the very phenomena that started off discussion. Others claim that the only coherent notion of consciousness we can have is one defined in terms of broadly intentional states, higher-order thoughts about what is going on, as we say, in our minds. It is just too early to say now whether all the detailed work in psychology and philosophy on these topics will come to be seen as advances, or just as variations on the traditional ways of taking sides.

So far I have considered three ways of responding to the explanatory gap. The first, which I only just mentioned, essentially consists in asking: what gap? Here the thought is that consciousness, as well for that matter as the intentionality of contentful states, are just states of the human brain. The next two notice the gap, and see it as a reason either to deny physicalism or to deny the coherence of the cartesian picture. In a way, these two positions are really variations of the 'What gap?' response, since, if either of them is accepted, the supposed mountains on the 'other

side' are found too small for there to be a gap. The fourth response takes the gap itself seriously, and insists that only a kind of arrogance could suggest that we will ever cross it. More concretely, the idea is that, conceptually bridging the gap between our intuitions about experiential phenomena and our picture of the world as at bottom physical, might well require a greater intelligence than is available to human beings. Our intellectual capacities are clearly not boundless, so why not count this particular problem as beyond our ken?

### *Action*

Actions are what we do rather than what merely happens to us. But not everything we do counts as an action. Snoring is certainly something done, but it seems not to be an action. Actions are tied, in some sense, to planning, or at least overall responsibility, and both are absent whilst one is asleep. In contrast, of a typical case of acting—say, putting the kettle on the stove—one feels one is responsible for it, and that one has set oneself to do it with some kind of forward planning.

Some might find it odd that I have classified actions as the third major category of mind. There is a tendency—perhaps due to the cartesianism discussed above—to think of the mind as fundamentally a locus of experiences. Attitudes are of course admitted, but the naive intuition is that attitudes are states we are aware of, and in this way are assimilable within the cartesian conception. But actions are typically regarded as the *products* of the mind, as movements of the body brought about by the experiences and attitudes of the minded subject. An image I find useful here is that actions are like the wake left behind as the mind moves through its environment. From this perspective, it can seem as if actions should not count as phenomena of

mind, on all fours with items in the other categories. This tendency is reinforced by fact that the theory of action has historically tended to be less integrated into the philosophy of mind than the study of attitudes and experience. On the one hand, it has figured more centrally in metaphysics and ethics than in philosophy of mind. And on the other, it has not seemed that actions are a threat to, or are threatened by, physicalism. However, all of this is changing and, I think it fair to say, the idea of actions as mere products of the mind has less of a grip, even if it remains an initial prejudice of the neophyte. (It has never been obvious to me why people are drawn to this view of actions in the first place. After all, many things we do are, as it were, 'internal' to the mind. Thus, deciding, calculating in one's head, perceiving and believing are all what one might call 'mental acts'; they are associated with no obvious bodily movements. Surely, we need an account of these in any reasonable philosophy of mind.)

There are many reasons for the change. First, the broadly anti-cartesian conception of the mind which one finds in Wittgenstein and others has led to a greater focus on action, and the role of the body in acting. Secondly, the recent search for an understanding of folk psychology has tended to bring a more sophisticated account of action in its train. The delicate task of saying what reasons are, and how they are related to causes of action, has tended to re-open the debate about the nature of action. Thirdly, as more work has been done on the metaphysics of action, it has become increasingly obvious that certain philosophical and commonsense intuitions about action are simply inadequate. We understand them less than we may once have thought, and the trail to a better understanding leads straight into other phenomena of mind. Finally, recent work in psychology on the initiation and planning of action has led philosophers to reconsider our commonsense intuitions.

The metaphysics of action begins with the question: what kind of thing is an action? As noted above, a first step is to say that actions are things we do, but this only puts off the deeper issues. For it is far from clear what kind of thing is ‘something done’. Progress seems to come from the recognition that actions belong to the broader class of events, but the hard work then consists in saying what events are, and in setting out the differentia of action within the class of events. A basic decision needs to be made at this point. Some writers regard events as objects consisting of the exemplification of a property at a time and place, and actions are then a sub-class of these events. On this view, my putting the kettle on the cooker and my boiling water are two different actions even though they take place at the same time and involve the same kettle and cooker. Others regard them as particulars, sharing this particularity with individual physical objects, despite the ways in which events and objects differ. On this view, I engage in only one action when I put the kettle on, but it can be described either as putting the kettle on or boiling water .

The debate between the property and particularity theorists can become quite complicated, and our ordinary ways of talking about what we do, play a large part in it. But the later is often less clear than philosophers would like, so it is not easy to decide in favour of either view. More significant, however, is the issue of the differentia of action; once it is admitted that they are events (whether of the property exemplifying kind or as particulars), we need to find out what makes some events actions. Taking a very synoptic view, there are two main contenders here. First, is the view that an action is some bodily movement which has an appropriate mentalistic description or explanation. Thus, if we can explain my hand’s moving only by saying that something pushed it, we do not regard me as having acted. But if its moving is describable as a case of inten-

tional reaching, and/or if I had a reason for reaching, the movement becomes an action.

This first view suggests that actions are intelligible as such within a mentalistic context, but it encourages the idea that at bottom actions are just bodily movements; we need to put them into a mentalistic context to see them, but what is put into such a context is not more special than any other event of movement. In contrast, the second view insists that an action is itself something with a mental 'part'. Whether the older terms 'will' or 'volition' are used, or the more contemporary 'trying' or 'setting oneself to', the idea here is that an action is a complex consisting of an event of volition and a bodily movement caused by that volition. What is a volition? Unlike historical attempts to describe it, the volition is not itself counted as an act, as this leads to the incoherence of regress. But it is counted an occurrence, event or phenomenon having something of the same status as bodily or perceptual sensations.

If actions partly consist in, and are not merely caused or explained by, phenomena of mind, the ordinary intuition that actions are mere products of the mind looks lame. Yet there are good reasons for accepting the now very popular volitional theory, both because it seems to give a plausible account of the dual nature of action and because it coheres with a lot of other things we think about the mind and its body.

### *Conclusion*

I have perhaps gone on long enough, but there is a sense in which a main part of the subject has only been touched on. For I have yet to say anything about how the different categories fit together in that repository of the mind generally known as the self. This is partly a reflection of the recent history of philosophy of mind. Up until the mid-70s, many



of the topics now central to the philosophy of mind were spread across the spectrum of philosophical sub-disciplines. Work on propositional attitudes was largely confined to philosophical logic and the philosophy of language. Philosophical discussions of action, including the issue of free-will, could be found in metaphysics, philosophy of science (especially social science), and in moral philosophy. And consciousness or experience figured in the epistemology of perception, as well of course as in the metaphysics of the mind-body question.

More recently, these different phenomena have come to be treated in more detail and in their own right. This was partly a result of the hopes for the naturalization of epistemology and even more because of philosophers looking over their shoulders at developments in psychology, neuroscience and computation. However, interest in the self, or personhood, continued to figure mostly in the metaphysics of personal identity, and there tended to be little direct concern with it as the focus of attitude, experience and action. I suppose there seemed no need to ask even such a basic question as whether the subject of belief, the sufferer of consciousness and the agent of action were a unity in any genuine sense. Perhaps, for the purposes of metaphysics and logic, it just seemed obvious that the answer was affirmative.

The only areas in which there has been some discussion of the self as a phenomena of mind have been in the philosophy of psychoanalysis and in the many attempts to describe self-deception. But I think that this is changing. The fragmentation of the mind is to a large extent an epiphenomenon of the way in which the philosophy of mind developed, and there are signs now of attempts to re-unify the field. To be sure, this has created a tension between those who think that the phenomena themselves only make sense as features of persons and those who continue to think that

progress can be made in a more piecemeal way. But insofar as this tension comes to generate a proper reasoned debate, we will see the philosophy of mind reconnected to the more traditional concerns of, among others, Hume and Kant.

## BIBLIOGRAPHY

### *General (Readers and Reference)*

The following are collections of articles central to the philosophy of mind and each includes detailed bibliographies. Guttenplan 1994 is an alphabetically arranged and in-depth survey of the whole subject.

Block, N (ed.), 1980, *Readings in the Philosophy of Psychology*, two vols., Harvard University Press, Cambridge (Mass.).

Goldman, A. (ed.), 1993, *Readings in Philosophy and Cognitive Science*, MIT Press, Cambridge (Mass.).

Guttenplan, S. (ed.), 1994, *The Companion to the Philosophy of Mind*, Basil Blackwell, Oxford.

Lycan, W.G. (ed.), 1990, *Mind and Cognition*, Basil Blackwell, Oxford.

### *General (Authors)*

Some of the main writers on the philosophy of mind are difficult to pigeonhole. Certain of those listed here have written about their own work in Guttenplan 1994 (above), but a list of central writings is:

Armstrong, D.M., 1968, *A Materialist Theory of Mind*, Humanities Press, London.

Davidson, D., 1980, *Essays on Action and Events*, Oxford University Press.

Dennett, D., 1978, *Brainstorms: Philosophical Essays on Mind and Psychology*, Montgomery, VT, Bradford.

—, 1987, *The Intentional Stance*, MIT Press, Cambridge (Mass.).

- , 1991, *Consciousness Explained*, Little Brown, Boston.
- Dretske, F., 1981, *Knowledge and the Flow of Information*, MIT Press, Cambridge (Mass.).
- , 1988, *Explaining Behaviour*, MIT Press, Cambridge (Mass.).
- Fodor, J., 1983, *The Modularity of Mind*, MIT Press, Cambridge (Mass.).
- , 1987, *Psychosemantics*, MIT Press, Cambridge (Mass.).
- , 1990, *A Theory of Content*, MIT Press, Cambridge (Mass.).
- , 1994, *The Elm and the Expert*, MIT Press, Cambridge (Mass.).
- Lewis, D., 1983, *Philosophical Papers*, vol. 1, Oxford University Press.
- , 1986, *Philosophical Papers*, vol. 2, Oxford University Press.
- Putnam, H., 1975, *Mind, Language and Reality: Philosophical Papers*, vol. 2, Cambridge University Press.
- , 1988, *Representation and Reality*, MIT Press, Cambridge (Mass.).
- Quine, W.V.O., 1960, *Word and Object*, MIT Press, Cambridge (Mass.).
- Ryle, G., 1949, *The Concept of Mind*, Hutchinson, London.
- Searle, J., 1983, *Intentionality*, Cambridge University Press.
- , 1992, *The Rediscovery of Mind*, MIT Press, Cambridge (Mass.).
- Stalnaker, R., 1984, *Inquiry*, MIT Press, Cambridge (Mass.).
- Wittgenstein, L., 1953, *Philosophical Investigations*, Translated by G.E.M. Anscombe, Basil Blackwell, Oxford.

### *The Semantics of Content*

- Crimmins, M., 1992, *Talk About Belief*, MIT Press, Cambridge (Mass.).
- Frege, G., 1966, *Translations from the Philosophical Writings*, edited by P. Geach and M. Black, Basil Blackwell, Oxford.
- Lycan, W.G., 1984, *Logical Form in Natural Language*, MIT Press, Cambridge (Mass.).
- Quine, W.V.O., 1960, *Word and Object*, MIT Press, Cambridge (Mass.).

- Richard, M., 1990, *Propositional Attitudes*, Cambridge University Press.
- Salmon, N., 1984, *Frege's Puzzle*, MIT Press, Cambridge (Mass.).
- Salmon, N. and Soames, S. (eds.), 1988, *Propositional Attitudes*, Oxford University Press. (This is a useful collection of readings in the area.)
- Schiffer, S., 1987, *Remnants of Meaning*, MIT Press, Cambridge (Mass.).

*The Nature of Content*

- Baker, L.R., 1987, *Saving Belief: A Critique of Physicalism*, Princeton University Press.
- , 1995, *Explaining Attitudes*, Cambridge University Press.
- Barwise, J. and Perry, J., 1983, *Situations and Attitudes*, MIT Press, Cambridge (Mass.).
- Burge, T., 1979, "Individualism and the Mental", *Midwest Studies in Philosophy*, 4, pp. 73–121.
- Cummins, R., 1989, *Meaning and Mental Representation*, MIT Press, Cambridge (Mass.).
- Davidson, D., 1984, "On Saying That", in his *Inquiries into Truth and Interpretation*, Oxford University Press.
- Dretske, F., 1981, *Knowledge and the Flow of Information*, MIT Press, Cambridge (Mass.).
- , 1988, *Explaining Behaviour*, MIT Press, Cambridge (Mass.).
- Field, H., 1978, "Mental Representation", *Erkenntnis*, 13, pp. 9–61.
- Fodor, J., 1981, *Representations*, Harvester Press, Hassocks.
- , 1987, *Psychosemantics*, MIT Press, Cambridge (Mass.).
- , 1990, *A Theory of Content*, MIT Press, Cambridge (Mass.).
- McDowell, J. and Pettit, P. (eds.), 1986, *Subject, Thought and Context*, Oxford University Press.
- McGinn, C., 1989, *Mental Content*, Basil Blackwell, Oxford.
- Millikan, R., 1984, *Language, Thought and Other Biological Categories*, MIT Press, Cambridge (Mass.).
- Papineau, D., 1993, *Philosophical Naturalism*, Basil Blackwell, Oxford.

- Peacocke, C., 1986, *Thoughts: An Essay on Content*, Basil Blackwell, Oxford.
- Putnam, H., 1988, *Representation and Reality*, MIT Press, Cambridge (Mass.).
- Searle, J., 1983, *Intentionality*, Cambridge University Press.
- Stalnaker, R., 1984, *Inquiry*, MIT Press, Cambridge (Mass.).
- Stich, S., 1992, "What is a Theory of Mental Representation?", *Mind*, 101, pp. 243–261.
- , 1983, *From Folk Psychology to Cognitive Science: The Case Against Belief*, MIT Press, Cambridge (Mass.).

### *Physicalism and Identity*

- Armstrong, D., 1968, *A Materialist Theory of Mind*, Humanities Press, London.
- Charles, D. and Lennon, K. (eds.), 1992, *Reduction, Explanation and Realism*, Oxford University Press. (A very good collection on this topic.)
- Churchland, P.M., 1981, "Eliminative Materialism and the Propositional Attitudes", *Journal of Philosophy*, 78, pp. 67–90.
- Crane, T. and Mellor, D.H., 1990, "There is no Question of Physicalism", *Mind*, 99, pp. 185–206.
- Kripke, S., 1980, *Naming and Necessity*, Harvard University Press, Cambridge (Mass.).
- Lewis, D., 1983, "An Argument for the Identity Theory", in his *Philosophical Papers*, vol. 1, Oxford University Press.
- , 1983, "Mad Pain and Martian Pain", in his *Philosophical Papers*, vol. 1, Oxford University Press.

### *Wittgenstein and the Private Language Argument*

- Budd, M., 1989, *Wittgenstein's Philosophy of Psychology*, Routledge and Kegan Paul, London.
- Kripke, S., 1982, *Wittgenstein on Rules and Private Language*, Harvard University Press, Cambridge (Mass.).
- McDowell, J., 1992, "Meaning and Intentionality in Wittgenstein's Later Philosophy", in *The Wittgenstein Legacy*,

edited by P. French, T. Uehling and H. Wettstein, University of Notre Dame Press, Notre Dame.

### *Emotion*

- Baier, A., 1990, "What Emotions Are About", *Philosophical Perspectives*, 4, pp. 1–29.
- de Sousa, R., 1987, *The Rationality of the Emotions*, MIT Press, Cambridge (Mass.).
- Gordon, R., 1987, *The Structure of Emotions*, Cambridge University Press.
- Nussbaum, M., 1990, *Love's Knowledge*, Oxford University Press.

### *Computational Models*

- Bechtel, W. and Abrahamson, A., 1991, *Connectionism and the Mind*, Basil Blackwell, Oxford.
- Clark, A., 1989, *Microcognition*, MIT Press, Cambridge (Mass.).
- Churchland, P.M., 1989, *A Neurocomputational Perspective. The Nature of Mind and the Structure of Science*, MIT Press.

### *Functionalism*

- Block, N. (ed.), 1980, *Readings in the Philosophy of Psychology*, vol. 1, Harvard University Press, Cambridge (Mass.).
- , 1978, *Troubles with Functionalism*. In *Perception and Cognition: Minnesota Studies in the Philosophy of Science*. Edited by W. Savage, University of Minnesota Press, Minneapolis.
- Loar, B., 1981, *Mind and Meaning*, Cambridge University Press.
- Putnam, H., 1975, *Mind, Language and Reality: Philosophical Papers*, vol. 2, Cambridge University Press.
- , 1988, *Representation and Reality*, MIT Press, Cambridge (Mass.).
- Shoemaker, S., 1981, "Some Varieties of Functionalism", *Philosophical Topics*, 12, pp. 93–120.

—, 1984, *Identity, Cause and Mind*, Cornell University Press, Ithaca.

See also: Rosenthal 1991, Lycan 1990, Guttenplan 1994 (in General above).

### *The Nature of Folk Psychology*

Churchland, P.M., 1984, *Matter and Consciousness*, MIT Press, Cambridge (Mass.).

Davies, M. and Stone, T. (eds.), 1995, *Folk Psychology*, Basil Blackwell, Oxford.

—, 1995, *Mental Simulation*, Basil Blackwell, Oxford.

The above two collections are very complete and include a great deal of further bibliography.

Goldman, A., 1992, "In Defense of Simulation Theory", *Mind and Language*, 7, pp. 104–119.

Gordon, R., 1986, "Folk Psychology as Simulation", *Mind and Language*, 1, pp. 158–171.

Greenwood, J., 1991, *The Future of Folk Psychology*, Cambridge University Press.

See also Baker 1995, Fodor 1987, Stalnaker 1984, Stich 1983 all listed in *The Nature of Content* above.

### *Philosophy of Psychoanalysis*

Cavell, M., 1993, *The Psychoanalytic Mind: From Freud to Philosophy*, Harvard University Press, Cambridge (Mass.)

Gardner, S., 1993, *Irrationality and the Philosophy of Psychoanalysis*, Cambridge University Press.

Wollheim, R. and Hopkins, J. (eds.), 1982, *Philosophical Essays on Freud*, Cambridge University Press.

### *Language of Thought*

Clark, A., 1989, *Microcognition*, MIT Press, Cambridge (Mass.).

Fodor, J., 1975, *The Language of Thought*, Thomas Y. Crowell, Nueva York.

—, 1987, *Psychosemantics*, MIT Press, Cambridge (Mass.).

Maloney, J., 1989, *The Mundane Matter of the Mental Language*, Cambridge University Press.

### *Supervenience and Reduction*

Davidson, D., 1980, "Mental Events", in his *Essays on Actions and Events*, Oxford University Press.

Horgan, T., 1982, "Supervenience and Microphysics", *Pacific Philosophical Quarterly*, 63, pp. 29–43.

Kim, J., 1993, *Supervenience and Mind*, Cambridge University Press.

### *Epiphenomenalism and the Inertness of Content*

Davidson, D., 1993, "Thinking Causes", in Heil and Mele (eds.). Heil, J. and Mele, A. (eds.), 1993, *Mental Causation*, Oxford University Press.

Jackson, F. and Pettit, P., 1988, "Broad Contents and Functionalism", *Mind*, 47, pp. 381–400.

Kim, J., 1984, "Epiphenomenal and Supervenient Causation", *Midwest Studies in Philosophy*, 9, pp. 257–270.

### *Experience and Consciousness*

Block, N., 1978, "Troubles with Functionalism", in *Perception and Cognition: Minnesota Studies in the Philosophy of Science*, edited by W. Savage, University of Minnesota Press, Minneapolis.

Davies, M. and Humphreys, G. (eds.), 1993, *Consciousness*, Basil Blackwell, Oxford. (This contains a very complete bibliography.)

Dennett, D., 1991, *Consciousness Explained*, Little Brown, Boston.

Flanagan, O., *Consciousness Reconsidered*, MIT Press, Cambridge (Mass.).

Jackson, F., 1982, "Epiphenomenal Qualia", *Philosophical Quarterly*, 32, pp. 127–136.

—, 1986, "What Mary didn't know", *Journal of Philosophy*, 83, pp. 291–295.



Lycan, W.G., 1987, *Consciousness*, MIT Press, Cambridge (Mass.).  
McGinn, C., 1983, *The Subjective View*, Oxford University Press.  
Nagel, T., 1974, "What is it like to be a bat?", *Philosophical Review*, 83, pp. 435–450.  
Rosenthal, D., 1986, "Two Concepts of Consciousness", *Philosophical Studies*, 49, pp. 329–359.  
See also: Rosenthal 1991, Lycan 1990, Guttenplan 1994 (in General above).

### *Action*

Bishop, J., 1989, *Natural Agency*, Cambridge University Press.  
Brand, M., 1984, *Intending and Acting*, MIT Press, Cambridge (Mass.).  
Bratman, M., 1987, *Intention, Plans and Practical Reason*, Harvard University Press, Cambridge (Mass.).  
Ginet, C., 1990, *On Action*, Cambridge University Press.  
Mele, A., 1992, *Springs of Action*, Oxford University Press.  
O'Shaughnessy, B., 1980, *The Will*, 2 vols., Cambridge University Press.  
See also Davidson 1980, Dretske 1988, in General (Authors) above.

*Recibido: 18 de marzo de 1996*