

OUP Copyright

CHAPTER 25

Reasons and the content of mental states: 2. Antireductionism and discursive psychology

Chapter contents

- Session 1 **The discursive alternative** 688
- Session 2 **Wittgensteinian approaches to mental content** 690
- Session 3 **Dennett and the Intentional Stance** 696
- Session 4 **Davidson and Radical Interpretation** 696
- Session 5 **Singular thought and the division between mind and world** 707
- Session 6 **Discursive psychology and Alzheimer's disease** 711
- Reading guide** 714
- References** 714

OUP Copyright

While the previous chapter focused on cognitivist approaches to intentionality...

Chapter 24 examined the kind of approach to meaning and intentionality implicit in cognitive neuropsychology and cognitive neuropsychiatry. The underlying approach was characterizable using a term from the philosophy of thought and language as representationalism. Such an approach aims to shed light on how our mental states can have the meaning or content they do using a cognitivist model drawing heavily on an analogy with computing. A key idea is that the fact that people can have mental states such as beliefs can be explained at a lower subpersonal level in terms of information processing using inner mental representations: structural configurations in their brains. Because the intentionality—the world-involving nature—of the mental states of whole people is explained in terms of the causal processes acting on lower level parts of their brains, such an approach can be thought of as ‘bottom up’.

... this chapter will explore an alternative non-reductionist perspective

This chapter will examine a contrasting view. In psychiatry this is sometimes called a ‘discursive approach’. In one of the readings of the chapter (linked with Exercise 1) Steven Sabat and Rom Harré (1994) characterize this as the view that ‘meanings are jointly constituted by the participants to a conversation’. The precise idea behind the phrase ‘jointly constituted’ will be subject to scrutiny, but whatever the details, if meanings are so constituted then they are not simply the result of causal processes going on in individual minds. Taking this talk of ‘constitution’ seriously implies that the kind of reductionist explanation of meaning, which is the aim of cognitivist approaches cannot succeed. Again, light will be shed on the general approach exemplified by discursive psychology by looking to some philosophical models that explore its underlying assumptions.

The philosophical alternative

In the philosophy of thought and language (also called the philosophy of content), an alternative to a cognitivist or representationalist theory of content has been developed by a number of different philosophers including Dennett, Davidson, Wittgenstein, McDowell, and the neo-Fregeans. These do not share a single, tightly defined, theory, but they do agree on a broad alternative strategy to representationalism, which could be called discursive. It might be characterized through four claims:

1. It is antireductionist, explaining intentionality not through internal brain states but as a way of acting in the world.
2. It grants an important role to the ‘interpretation’ of speech and action from an everyday third-person perspective.
3. As a result of points 1 and 2 it is called an ‘externalist’ approach.
4. There is a central connection between our intentionality and *rationality*.

These points will now be briefly summarized.

Antireductionism and the practical turn

The alternative to the philosophical project of naturalizing meaning or, more broadly, ‘intentionality’ through a reductionist account such as representationalism, is to provide a non-reductive account based on practice. Instead of construing thought as a system of internal representations, the alternative is to construe it as a systematic practical engagement with the world. Whereas representationalism or cognitivism aims to explain intentionality in meaning-free terms using causal relations between inner and outer states, the authors listed above share the view that intentionality cannot be broken down like that. It has instead to be characterized in meaning-presupposing terms and pertaining to the behaviour of whole people not parts of their brains.

Meaning is necessarily available to a third person stance

In addition, Davidson, Dennett, and to some extent Wittgenstein and McDowell argue that an understanding of both thought and language requires understanding how behaviour is understandable from a mundane third person perspective. Light is shed on intentionality by examining the explanatory stance that we can take to other people to make sense of their speech and action. Thus Davidson (1984a) approaches intentionality by asking what the preconditions are for interpretation from scratch in ‘Radical interpretation’ while Dennett discusses the underpinnings of the ‘Intentional Stance’.

An aside on the connection with epistemology in Chapter 27

Given this connection between intentionality and the third person perspective, there will be some connections between this chapter and Chapter 27 on autism and our knowledge of other minds. A key recent debate in the epistemological of mind has been between two opposing views of what our knowledge of other people’s mental states is based. One approach, the ‘theory-theory’, is based on the idea that in everyday circumstances we apply a tacit *theory* of how minds in general work. The other approach, ‘simulation theory’, argues instead that we imaginatively reproduce what other people are thinking by seeing what we would think in those circumstances. On this latter view you do not need to have a *theory* of minds to know what other people are thinking just a *mind* yourself. In fact, although the authors that will be considered in this chapter emphasize the role of the third person perspective they do not all subscribe to a theory-theory. That is because the subject of this chapter is not epistemology but rather the ontology of minds. This point will become clearer in Chapter 27, however.

Externalism

One consequence of stressing the theoretical role of the third person stance in accounting for intentionality is externalism. Both internalists and externalists can agree that whether a belief is *true* or not depends on features of the external world (assuming that it

is a belief about some feature of the external world). But externalists, unlike internalists, argue that at least in the case of some beliefs, their very *content* depends on relations to the world. Thus, for example, they dispute Descartes' assumption in the *Meditations* that the content of one's thoughts could be just the same even if there were no external world. Descartes held that in the absence of the external world, many of one's beliefs would turn out to be false. Externalists hold that one could not even entertain many of the beliefs we do.

Rationality

Finally, an important element of thinking about the content of mental states in the antireductionist way discussed below is the central role of rationality. A key idea is that we adopt an interpretative or explanatory stance to other people, which differs from the stance appropriate to non-human objects such as chairs and tables (animals have an interesting intermediate status). In the case of humans (and some animals) we make *sense* of their behaviour (and we might want to say *action* rather than *behaviour* to emphasize this point). But making sense of other people requires, as a constitute principle, that are able to find their behaviour broadly speaking *rational*. Thus rationality is of central importance in this chapter (by contrast with its modest role in Chapter 24). It is a precondition of being 'minded' that one is rational.

Plan of the chapter

- ◆ *Session 1* outlines in general terms the sort of alternative approach that will be discussed in this chapter.
- ◆ *Session 2* will examine three philosophical interpretations of the work of the later Wittgenstein. The aim will be both to sharpen the criticism of representationalist or cognitivist approaches discussed in Chapter 24 and to consider how best to understand Wittgenstein's alternative view of the mind and mental content.
- ◆ *Session 3* will examine the influential work of the American philosopher of mind Daniel Dennett.
- ◆ *Session 4* will examine the related but different view developed by Donald Davidson.
- ◆ *Session 5* will examine the key assumptions behind neo-Fregean approaches to thought.
- ◆ *Session 6* will return to the clinical reading from Session 1 on a Discursive Psychological approach to Alzheimer's disease and ask what general conclusions can be drawn.

Session 1 The discursive alternative

This short session will set the scene for the rest of the chapter. The extract linked with Exercise 1 below is just the first part of a paper, to which we will return in the final session (see reading linked with Exercise 8), by the psychologist Steven Sabat and the philosopher Rom Harré on Alzheimer's disease sufferers. However, this first part of the paper concerns the broader underpinnings of their thinking about meaning.

EXERCISE 1

(15 minutes)

Read:

Sabat, S.R. and Harré, R. (1994). The Alzheimer's disease sufferer as a semiotic subject. *Philosophy, Psychiatry, & Psychology*, 1: 145–160. (Extract: pp 145–148)

Link with Reading 25.1

- ◆ What is the view of meaning to which the authors subscribe?

The view of meaning

There are a number of passages in which Sabat and Harré reveal their view of meaning and of the intentional nature of mental states in general. (Remember that in the context of the philosophy of thought and language, which is sometimes called the philosophy of content, the term 'intentionality' generally means the *aboutness* or *world-involving* nature of both mental states and linguistic utterances. This term was reintroduced to philosophy by Franz Brentano (1838–1917) a philosopher and psychologist. The general characteristics were discussed at some length in Chapter 24.)

The abstract to the paper comments that 'meanings are jointly constituted by the participants to a conversation'. However, also 'from a discursive point of view, psychological phenomena are not inner or hidden properties or processes of mind which discourse merely expresses. The discursive expression is, with some exceptions, the psychological phenomenon itself' (p. 146). They recruit Leo Vygotsky (1896–1934), the Soviet psychologist, in support of the thesis that: "The mind" is no more than, but no less than, a privatised part of the "general conversation". Meanings are jointly constructed by competent actors in the course of projects that are realized within systems of public norms' (p. 146).

More precisely the paper reports that it uses the word 'meaning' in three different ways (p. 147):

1. for intention in acting,
2. for an interpretation of events and situations, and
3. for evaluations of events, situations, or actions.

Why are these three all regarded as senses of the same concept? Well in each case there is an idea of *significance* at work. Starting with the second point, the idea of interpreting events is simply that of seeing what the events are about or perhaps more broadly, just what the events are, such as, for example, a robbery, a wedding, a promise, etc. To construe the collection of sounds and movements involved as making up any of these actions is to fit them into a broader sense-giving context. (Because of what it is, a wedding ceremony has particular consequences that differ from those of a robbery.)

Once so construed, these actions and events can be evaluated in a variety of ways (point 3 above). These include the assessment of their desirability in the selfish idea of the wants and desires of an agent but also in the distinct sense of their rationality in the light

of the agents other beliefs and desires. Although one may not approve of someone's voting intentions, one may be able to see how it is rational in the light of their other beliefs and about the world and their values. Finally (to take the first point), an agent can then act in response to or to bring about just such events, again because of the significance they have for him or her. What is more, an action is the action it is in part in virtue of the agents reasons for acting. (Administering a drug may ease a patient's pain and also hasten her death, but if the reason for administering the drug was the former and not the latter then the action is an action of easing pain rather than a killing.)

The contrast of emphasis with Chapter 24

Without going further into the paper, it is possible to see the sharp contrast in *emphasis* between this approach to meaning and the view discussed in Chapter 24. It is one thing to talk about the meaning of events, and to think of the actions of individuals in terms of events (bringing them about, preventing them, etc.) and quite another to talk of the meaning or content of psychological states in terms of second order structures within the brain, or computations over representations in an internal 'semantic system'.

But while there is a contrast in emphasis, is there really a clash in the underlying claims that these apparently different views take on meaning? Or might it be simply that they emphasize two different aspects of meaning that could, in fact, be reconciled in a future science of the mind? If meaning really is constituted jointly, how is this brought about? And what relation does this have to do with what goes on in the brain? Again to be clear about the underlying ideas it is useful to turn to contrasting philosophical models. These will help clarify the extent to which there might or might not be a reconciliation between the different perspectives on meaning and mental content.

A challenge to reductionists . . .

But to set the scene, think back to a challenge that was raised against reductionism in Chapter 24 and to which we will return towards the end of this chapter. The challenge is that thinking about meaning from the perspective of cognitivism arguably makes a mystery of the connection between states of mind and states of the world. As soon as one talks of internal mental representations meaning second order structures of the brain then it becomes mysterious how these have any bearing on the world. So if one wants to explain how it is possible for me to have thoughts about the cup in front of me, as soon as one postulates a state of my brain and suggests that it might be a representation of that cup, the question of how that configuration of neurones, for example, can have anything to do with the cup becomes pressing. Why isn't it *just* a configuration of neurones?

. . . and a reciprocal challenge to antireductionists

Nevertheless, there remains something problematic about the idea of intentionality to which the reductionists are, at least, sensitive. This was well captured by the quotation from Fodor in the previous chapter. As 'aboutness' is not likely to be one the explanatory

concepts used in a complete physics, then if it is a genuine feature of the world, it seems that it must really be or comprise something else. How else can something as mysterious as intentionality be part of the physical universe, on the assumption that we are not embracing an antiscientific dualism of mind and matter? This in turn raises a challenge to antireductionists. How can they make intentionality an unmysterious part of the world?

How can antireductionists fit meaning into nature?

As we saw in Chapter 24, one approach is to attempt a *reductionist* explanation of intentionality by devising causal and or evolutionary mechanisms to explain how internal states can take on or encode mental content. Such an approach has the advantage in that it aims to explain the problematic notions of content, meaning, and intentionality in supposedly better understood notions such as causation. But the alternative, which will be discussed in this chapter, takes a different tack. To anticipate, it aims to make these notions less mysterious by thinking instead about the everyday role of ascribing content-laden mental states to people. Thus both Dennett and Davidson suggest that the idea of content-laden mental states have to be understood as playing a role in the explanation and prediction of people's speech and action. They play a role within the overall interpretative strategy of either what Dennett calls the 'Intentional Stance' or what Davidson calls 'Radical Interpretation'. These will be the subject of Sessions 3 and 4 below.

Neither Dennett nor Davidson attempt to *reduce* the ideas of meaning, content, and intentionality to less problematic concepts. Neither the Intentional Stance nor Radical Interpretation can be characterized in merely causal terms, for example. They are instead approaches that can only be described as ways of finding patterns of meaning and significance in speech and action. Furthermore, as we will see, *rationality* plays an important central role in both accounts. So both approaches attempt to clarify connections between all these concepts but without showing how they might be reduced to the concepts of physical or even biological science.

Antireductionism and the discursive turn

How does this second general approach relate the comments in the reading above linked with Exercise 1 (Sabat and Harré, 1994, pp. 145–149) that meaning is constituted in interpersonal discourse? To a first approximation the answer is this. The philosophical view just sketched provides one way to unpack such a claim about how meaning is constituted. This is via the claim that it is a necessary feature of our speaking a meaningful language and having content-laden mental states, that our speech and action is interpretable by others in accordance with shared canons of rationality. This latter claim is one way of insisting that meanings (construed broadly to include the meaning of action and the 'meanings' of our beliefs and desires) are public. They are necessarily the sorts of things that can be articulated from the everyday third person perspective we take when we explain people's actions.

In fact, stronger claims about the social constitution of meaning are sometimes made. One such claim is that the content of our minds and utterances are not just necessarily available to others but that they are actually constructed piecemeal in social

negotiations. This sort of claim is a much more radical form of social constructionism. The weaker claim of the previous paragraph is consistent with the idea that when now I form the intention to end this sentence in an even number of words, that intention prescribes the range of satisfactory outcomes independently of any further social negotiation. It may be that it is necessarily the case that I could have been interpreted to have this intention (on the basis of what I say and do). But it is not necessary that I actually am so interpreted or that I speak to anyone about it to have the intention. Radical social constructionism by contrast insists that the connection between interpretation and mental states turns on what actually happens not on what could have happened. It will be touched on only rather briefly, however, because it is an implausible view.

Having now thought about the kind of alternative that might be available, the next session will turn to the later Wittgenstein's work via three different interpretations.

Reflection on the session and self-test questions

Write down your own reflections on the materials in this session drawing out any points that are particularly significant for you. Then write brief notes about the following:

1. What are the features of a discursive approach that relate specifically to meaning? What does it take meaning to be?
2. What prima facie advantages does this approach have over a cognitivist approach?
3. But is the approach really distinct from and incompatible with a cognitivist approach as described in Chapter 24?

Session 2 Wittgensteinian approaches to mental content

In order to clarify the idea of a social constructionist approach to meaning we will now turn to three interpretations of the work of the later Wittgenstein (1953) and his *Philosophical Investigations*. Wittgenstein's writing style is unusual: conversational German translated into conversational English with a minimum of technical terms. Nevertheless, it is not always easy to understand the general thread of his argument and thus there a number of rival interpretations.

One key idea, though, is that he stresses the public nature of meaning. But getting clear on what this entails is a matter for some work.

This session will outline some recent work on the later Wittgenstein's view of minds and meaning: by Saul Kripke, Crispin Wright, and John McDowell. The key positive idea is that understanding meaning should be thought of as a practical ability. This view is coupled with criticisms of the alternative discussed in Chapter 24: that having a meaning in mind is a matter of having an internal representation. To reach the positive account, however, it will be necessary to detour through that negative argument. The

best place to begin is with Kripke's notorious sceptical interpretation of Wittgenstein's arguments.

Saul Kripke's interpretation of Wittgenstein

Kripke's sceptical interpretation of Wittgenstein

Saul Kripke, known primarily for work on modal logic and the semantics of names, wrote a short interpretation of a key part of Wittgenstein's *Philosophical Investigations* in the 1970s called *Wittgenstein on Rules and Private Language* (1982). This has proved highly influential, not because it has commanded agreement with its interpretation, but because it has stimulated philosophers to diagnose just how it can apparently justify a conclusion that is so obviously wrong. Like many sceptical arguments, the conclusion is clearly disastrous, but on first inspection, the argument in support of it appears to be sound. What is instructive for the philosophy of content is to determine which assumption leads to the sceptical conclusion. The moral that will be drawn here is that the possession of content-laden mental states depends at root on practical abilities rather than free-standing internal mental representations or symbols.

With that final destination in mind, we will concentrate first on the force of Kripke's argument about addition. We will return shortly to its application to meaning and content more generally.

Adding and quadding, plus and quus

Kripke's challenge is based on a simple thought experiment. How could you justify that replying '125' is the right answer to the question 'what is $57 + 68$?'. Of course the first thing to do would be to ensure that there had been no mistake in carrying out the addition. One might best resolve this by setting out the addition using rows and columns, adding the units, correctly attending to the carry over and then adding the tens. But Kripke points out there is in principle a further 'meta-linguistic' error to be avoided that is giving the wrong interpretation to the sign '+'. He postulates a sceptic who proposes that in the past what you have meant by '+' was not the addition function but the 'quaddition' function which is defined thus:

$$x \text{ quus } y = x \text{ plus } y \text{ except where } x = 57 \text{ and } y = 68 \\ \text{where } 57 \text{ quus } 68 = 5$$

The sceptic argues that as this has been the function you have been carrying out in the past, the correct response to the question: 'What is $57 + 68$? is '5'. The question becomes one of establishing what one meant in the past when one used that sign. How can one tell what function one has previously used that sign to represent? How does one know that in the past one meant addition by it rather than, in Kripke's example, the 'bent function' 'quaddition'?

EXERCISE 2

(10 minutes)

Think what answer you could give to Kripke's sceptic. What answer could you give show that you are right not to answer 5 to what is $57 + 68$? What facts about your past performance or mental states show that you have meant addition rather than quaddition by '+' or 'plus'?

Adding and quadding, plus and quus again

Kripke's basic challenge stems from the fact that one's past behaviour is finite so one has only ever carried out a finite number of arithmetical operations. Thus the sceptical challenge is to say why writing '125' today is going on in the *same way* as one's past practice rather than, say, writing '5'. Sameness and difference depends on the rule or function that was previously applied. If previously one meant quaddition by '+' then writing '5' just is going on in the same way.

Kripke's two sceptical weapons

Kripke then deploys two sorts of consideration to undermine the efficacy of the most natural answers one might give to the sceptic. One device is the use of similar and ramifying (mis)interpretation of signs. The other is to stress the normativity implicit in the question. A satisfactory answer to the sceptic must show 'whether there is any fact that I meant plus not quus' (p. 11). But it must do this in such a way that it can 'show how I am *justified* in giving the answer 125' (p. 11). We will return to the role of this second element shortly.

The first tactic is evident in Kripke's responses to the following suggestions. In answer to the question, 'How does one now know which rule one was following in the past?', one might cite explicit instructions that one gave oneself. Perhaps one said (aloud): 'Now I'll *add* these numbers'. Obviously, this will not work because it simply repeats the problem. It only pins down one's past interpretation of the '+' sign if one can pin down the meaning of the spoken word 'add'. This problem applies equally whether the word is spoken aloud or silently to oneself.

The appeal to counting will not escape the problem of and quounting

One might hope instead to pin down the meaning of the sign '+' or the word 'add' by defining it—at some stage in the past—in more primitive counting terms. To add two numbers one counts on along the series of integers starting just after the first number and proceeding along by as many numbers as the second number. One might have repeated these instructions to oneself on some previous occasion. However, again this will not do because it will comprise a sequence of words that has correctly to be interpreted. Crucially as a way of distinguishing adding from quadding it depends what 'count' is taken to mean. If that really means *quount*—where to quount is the same as counting except where the numbers concerned are 57 and 68—the definition will serve to pick out quadding and not adding (Kripke, 1982, pp. 15–16).

It seems that any response to Kripkean scepticism that deploys any sort of sign or symbol that is spoken, written down or entertained in the mind will not work. Any sign could be misinterpreted to sustain a 'bent' rule.

Dispositionalism fails to address the normativity of the challenge

The other response, which Kripke considers and rejects, attempts to answer the sceptic by agreeing that meaning something by a

sign cannot consist in having any occurrent mental phenomena but arguing that it is instead a dispositional state. There are a number of problems with this response but the central objection is this. Dispositions cannot meet the second requirement highlighted above. They cannot by themselves *justify* an answer as correct because one may be *disposed* to make an error.

The dispositionalist gives a *descriptive* account of this relation: if '+' meant addition, then I will answer '125'. But this is not the proper account of the relation, which is *normative*, not descriptive. The point is *not* that, if I meant addition by '+', I *will* answer '125' but rather that, if I intend to accord with my past meaning of '+' I *should* answer '125'. (pp. 160–161)

Kripke also makes a further—less decisive—objection to dispositional accounts. This is that one's dispositions are also finite. Thus given sufficiently large numbers to add, one is not disposed to give the correct answer because, for example, one cannot accurately remember the numbers or add them in the head. Thus one's actual dispositions do not fix the correct interpretation of the '+' sign.

Kripke's sceptical conclusion

Kripke goes on to draw the following substantial conclusion. Given that no fact can be called to mind to determine which function one previously meant by '+' or 'addition', there are no such facts. Furthermore, the same arguments could subsequently be deployed for one's present use of signs. Furthermore, nothing turns on the mathematical nature of the example chosen. The same argument applies to the use of any word and its meaning and to any speaker. As the meaning of any word must turn on what speakers mean by it, he concludes: 'There can be no such thing as meaning anything by any word. Each new application we make is a leap in the dark; any present intention could be interpreted so as to accord with anything we may choose to do. So there can be neither accord, nor conflict.' (p. 55).

Epistemology and ontology

The fact that he draws a metaphysical conclusion from merely epistemological considerations may seem startling. How can Kripke draw such a conclusion about the link between rules and applications from epistemological considerations? The answer is that an important assumption is built into the sceptical approach. If there were some fact that constituted the relation between a rule and its applications, it would be independently identifiable by the idealized subject that Kripke postulates. Kripke supposes, for the purpose of argument, that one may have all the possible information about one's past experiences, mental states, and inclinations. He then asks whether any of these would be sufficient to determine the rule that one were following. His conclusion, based on his interpretation of Wittgenstein's arguments, is that none would be. Given the idealizations involved, and the assumption that had any fact constituted the rule one were following one would have known it, then there is no such fact of the matter.

This sceptical interpretation of Wittgenstein is reinforced by Kripke's reading of Wittgenstein's *Philosophical Investigation* §201. Wittgenstein writes there:

This was our paradox: no course of action could be determined by a rule, because every course of action can be made out to accord with the rule. The answer was: if everything can be made out to accord with the rule, then it can also be made out to conflict with it. And so there would be neither accord nor conflict here. Wittgenstein (1953, §201)

Thus Kripke denies that, properly speaking, there are any facts about meaning. What a word means is never itself a matter of truth or falsity.

A positive account of meaning and social constructionism

Given that sceptical conclusion, what positive account does Kripke give of meaning? He suggests that while an individual cannot be thought of as following a rule in isolation, the individual can be treated as following a rule in the context of a community. As long as they do not conflict with the community's judgements then they can be so regarded. The key idea here is not that facts about meaning are genuinely ushered back onto the stage. Rather, individuals can be *dignified* as rule followers providing that they do not conflict with the wider community.

It is essential to our concept of a rule that we maintain some such conditional as 'If Jones means addition by "+", then if he is asked for "68 + 57", he will reply "125"'. . . [T]he conditional as stated makes it appear that some mental state obtains in Jones that guarantees his performance of particular additions such as '68 + 57'—just what the sceptical argument denies. Wittgenstein's picture of the true situation concentrates on the contrapositive, and on justification conditions. If Jones does *not* come out with '125' when asked about '68 + 57', we cannot assert that he means addition by '+'. Kripke (1982, pp. 94–95)

This view can be taken to support a social constructionist view of meaning. While the correct application of a word can be specified in advance, something akin to correct use emerges from communal practice.

Mental content

It should not take too much thought to realize how the sceptical argument might be generalized to apply not only to meaning something by a word but also to other intentional mental states. Meaning something by a word is a matter of intending to use it in a particular way. One can ask more generally by what token one knows the content or intentional object of past mental states. And again, it seems that the result will be the sceptical destruction of our everyday understanding of mental states. Anything that came to mind and that might be supposed to have fixed the content of what one was thinking could be interpreted in any number of ways. Thus there will be no facts of the matter about a person's content-laden states of mind.

What has gone wrong to lead to this drastic conclusion?

The main moral of Kripke for cognitive approaches to psychiatry

It is worth being clear what the purpose of looking at Kripke has been for this chapter. Aside from whatever intrinsic interest there is in Kripke's argument, one conclusion will remain intact despite the criticisms of Kripke described below. If one attempts to construe mental states as inner representations, then one will fall to Kripke's sceptical argument. That is a lasting consequence of the passages of Wittgenstein on which it is based. Kripke's own conclusion—that there is no such thing as meaning—can be resisted without this showing that nothing important has been learnt through his presentation of Wittgenstein's destructive arguments. Where Kripke goes wrong is in ignoring the conditional nature of the claim: *if* one construes mental states as inner representations. . .

We will now turn to two philosophers who have diagnosed what is wrong with Kripke's argument.

EXERCISE 3

(10 minutes)

Think about Kripke's sceptical argument. Especially keeping the analogy between following a rule or understanding a meaning on the one hand and having an intentional state on the other, can any argument be offered against it?

Defusing Kripke's argument: Wright and McDowell on Wittgenstein

Kripke himself distinguishes between two sorts of response to a sceptical argument. There are sceptical responses, which accept the force of the sceptical argument but suggest a way in which its consequences can be 'lived with'. Kripke suggests that Hume's response to his own argument about causation is such. On the other hand there are 'straight' responses. These diagnose a fault in the sceptical argument. We will now discuss two such responses that aim to diagnose a fundamental flaw in Kripke's argument.

Crispin Wright's interpretation of Wittgenstein

Wright's diagnosis of Kripke's argument

In 'On making up one's mind: Wittgenstein on intention' Crispin Wright (1987) points out that the assumption that guides the various responses to the sceptic that Kripke considers and rejects is that the relevant epistemology is *inferential*.

Except for the dispositional response, all the answers take the form of postulating internal signs from which the meaning that one attached to '+' or 'addition' is supposed to be *inferred*. Even dispositions are deployed as the source of an inference about meaning. The sceptical conclusion results from the fact that no such unique meaning can be determined by these resources.

The alternative to an inferential epistemology is instead a form of direct memory of a *sui generis* state. Now it may seem that such direct access to a state cannot be possible because of the peculiar properties of meaning. How can one have access to a state that normatively prescribes the correct use of a word? However,



Fig 25.1 Crispin Wright

Wright points out that *intentions* have similar a property in that they *prescribe* the actions that satisfy them while we think of access to them as non-inferential.

Had the sceptical argument been directed against intention in general, rather than at what it is tempting to regard as the special case of meaning, there is no doubt that the intuitive concept seems to contain the resources for a direct rebuttal. Since I can know of my present intentions non-inferentially, it is not question-begging to respond to the Sceptic's challenge to my knowledge of my past intentions to reply that I may simply remember them. (p. 395)

Thus Wright suggests that Kripke's sceptical argument can be defused by pointing out that it rests on the assumption that access is inferential and that this is an implausible assumption given our everyday notion of intentions more broadly. But he agrees with Kripke that there is a substantial philosophical question lurking here that stands in need of an answer. (In this he disagrees with John McDowell as we will see below.)

Wright's own version of Kripke's problem

The problem that Wright (1987) sees is this. The fact that intending serves as a good analogy for understanding rules or meanings cuts in two directions. The analogy can be used to block Kripke's sceptic but it invites a closely related question to that which motivates the rule-following considerations. How does an intention that can be arrived at in a flash normatively constrain those actions that would accord with it in the future? As we have emphasized, the normative connection between an intention and what accords with it seems as mysterious as that between understanding a rule and its correct applications. Wright, in other words, re-emphasizes the fundamental connection between the

problem of linguistic meaning and mental content. Wright deploys constructivism to explicate both.

In another paper Wright says:

One of the most basic philosophical puzzles about intentional states is that they seem to straddle two conflicting paradigms: on the one hand they are avowable, so to that extent conform to the paradigm of sensation and other 'observable' phenomena of consciousness; on the other they answer constitutively to the ways in which the subject manifests them, and to that extent conform to the paradigm of psychological characteristics which, like irritability or modesty, are properly conceived as dispositional... It seems that neither an epistemology of observation—of pure introspection—nor one of inference can be harmonised with all aspects of the intentional. (p. 142)

Intention is only one example of a general phenomenon that also includes understanding, remembering, and deciding. In each case, the subject has a special non-inferential authority in ascribing these to herself which is, nevertheless, defeasible in the light of subsequent performance. Wright suggests that Wittgenstein's attack on reductionist explanations of such states shows that they cannot be modelled on a Cartesian picture of observation of private experiences. That is a significant result of even Kripke's summary of Wittgenstein's negative arguments. But if understanding, intending, and the like are to be modelled on abilities instead, as Wittgenstein seems to suggest but Kripke misses, how can the subject have special authority in ascribing these to herself in the light of the attack on reductionist explanation?

Wright's solution: constructivism

Wright's *constructivism* appears to provide a solution to this problem. The basic idea is to deny that there is any inner *epistemology* and to devise a constructivist account of intention instead:

The authority which our self-ascriptions of meaning, intention, and decision assume is not based on any kind of cognitive advantage, expertise or achievement. Rather it is, as it were, a *concession*, unofficially granted to anyone whom one takes seriously as a rational subject. It is, so to speak, such a subject's right to declare what he intends, what he intended, and what satisfies his intentions; and his possession of this right consists in the conferral upon such declarations, other things being equal, of a *constitutive* rather than descriptive role. (p. 400)

All other things being equal, a speaker's sincere judgements constitute the content of the intention, understanding, or decision. They determine, rather than reflect, the content of the state concerned. This sort of approach fits well with the initial summary of a discursive psychological approach at the start of this chapter in which meanings are constructed in ongoing conversations rather than fixed in the head by mental representations. Meanings and mental states are constructed in ongoing conversations.

Wright's account of first person access is of independent interest whether or not it reflects Wittgenstein's concerns. But for now it is important to note that it assumes that there is a real problem to be overcome. It assumes that meanings are not the sort of

thing that can be grasped without some further underlying story. To see how this assumption is not compelling, compare Wright's account of Kripke with that of McDowell.

John McDowell's interpretation of Wittgenstein

McDowell (1992) offers an initially similar diagnosis to Wright of the misleading assumption that underpins Kripke's argument. According to him, Kripke subscribes to a 'master thesis' about what can come before the mind: 'the thesis that whatever a person has in her mind, it is only by virtue of being interpreted in one of various possible ways that it can impose a sorting of extra-mental items into those that accord with it and those that do not' (p. 45).

Understanding meanings and mental states

This talk of 'sorting' is meant to capture the *normativity* implicit in understanding meaning. If one understands the meaning of a word one can sort the correct applications of it from the incorrect ones. One knows to which worldly items it applies. As McDowell (1992) emphasizes, this applies not only to understanding meaning but to intentionality in general. 'An intention, just as such, is something with which only certain future states of affairs would accord. Quite generally, a thought, just as such, is something with which only certain states of affairs would accord.' (pp. 45–46).

McDowell on Kripke's key assumption

Because Kripke assumes that the mind can only be populated by mental items that impose a normative constraint on worldly items once they have been interpreted, his sceptical deployment of a regress of interpretations is made possible. In other words, Kripke's assumption that the epistemology at play is based upon an inference follows from his view of what can come before the mind. Thus responses to the sceptic seem to have to attempt to justify knowledge of which rule guided past behaviour by inference from some free-standing mental item or state that could be interpreted in a number of different ways. Similarly access to one's current intentional states will require the interpretation of a mental item. 'The master thesis implies that whatever I have in my mind on this occasion, it cannot be something to whose very identity that normative link to the objective world is essential. It is at most something which *can* be interpreted in a way that introduces that normative link, although it can be interpreted differently.' (p. 46).

Scepticism defused

If the assumption that underlies the sceptical argument is that mental states require interpretation to connect them to actions or to the worldly items that satisfy them, then there is a way of escaping scepticism about meaning and intentionality. Instead of concluding from the fact that the normativity of meaning cannot be recovered from mental items that 'just stand there like a sign-post' that there are no facts about meaning, one can instead

conclude that it cannot be *reduced* to non-normative phenomena. The alternative is, in other words, to claim that meaning something by a word is a *sui generis* state. The fact that Kripke rejects this alternative suggests that he subscribes to an assumption like Fodor's representationalist claim that if meaning is real it must really be something else.

The only argument that Kripke has against non-reductionism about content is an argument from 'queerness' similar to that of Fodor:

Perhaps we may try to recoup, by arguing that meaning addition by 'plus' is a state even more *sui generis* than we have argued before. Perhaps it is simply a primitive state, not to be assimilated to sensations or headaches or any 'qualitative' states, nor to be assimilated to dispositions, but a state of a unique kind of its own... Such a state would have to be a finite object, contained in our finite minds. It does not consist in my explicitly thinking of each case of the addition table, nor even of my encoding each separate case in the brain: we lack the capacity for that. Yet (§195) 'in a *queer* way' each such case already is 'in some sense present'... What can that sense be? Can we conceive of a finite state which *could* not be interpreted in a quus-like way? How could that be?
Kripke (1982, pp. 51–52)

Kripke argues that any *sui generis* conception of meaning is too strange to form a natural part of the world. But in the passage just quoted he builds an important qualification into his description. A *sui generis* meaning would be a finite *object* contained in our finite minds. If this were the only model—meaning as a free-standing inner object—then Kripke would be right to reject a *sui generis* non-reducible conception of meaning. No object before the mind's eye could fix the normative consequences of meaning. But to assume that meaning must be conceived this way is already to be partially committed to the cognitivist or representationalist model. Effectively Kripke assumes that either a representationalist explanation of meaning is possible or, properly speaking, there is no such thing as meaning.

A hint towards a positive account

This diagnosis shows how scepticism can be avoided but it only gestures towards what a positive account of intentionality would be like. Wittgenstein provides one further clue in a passage the first half of which Kripke quotes. The passage continues:

It can be seen that there is a misunderstanding here from the mere fact that in the course of our argument we give one interpretation after another; as if each one contented us at least for a moment, until we thought of yet another standing behind it. What this shews is that there is a way of grasping a rule which is *not* an *interpretation*, but which is exhibited in what we call 'obeying the rule' and 'going against it' in actual cases.
Wittgenstein (1953, §201)

This passage begins to suggest a practical reorientation of the philosophy of content. Instead of thinking of mental states

as free-standing internal states or mental representations, one should think of them as more like practical abilities. But how like an ability can a state like an expectation really be? How can such talk shed light on intentionality? Further light on this question will be shed in the rest of this chapter. But a clue can be found by noting what is at issue between McDowell and Wright.

An overview of Wright and McDowell

Wright thinks that Kripke (and perhaps Wittgenstein) raises an interesting question about how we can have access to normative and prescriptive mental states. How can there be such states given that we can have immediate if fallible access to them? Wright then sets about a philosophical building project in which the content of intentional states—and thus also linguistic meaning—is constructed from subsequent judgements.

McDowell, by contrast, blames the need for this philosophical construction on a restricted picture of the sort of mental items that can come before the mind. He says:

The question 'How is it possible for meaning to reach ahead of any actual performance?' is just a specific form of the question 'How is it possible for the concept of accord to be in place in the way that the idea of meaning requires it to be?' The Wittgensteinian response is not that these are good questions, calling for constructive philosophy to answer them. The Wittgensteinian response is to draw attention to a defect in the way of thinking that makes it look as if there are problems here. (p. 49)

There are two important conclusions to draw:

1. Wright and McDowell agree in broad outline on the diagnosis of what is question begging about Kripkean scepticism. There is no need to *infer* what one meant on previous occasions by one's words from evidence that can be characterized in non-intentional terms. The attempt to explain the intentionality of mental states by postulating free-standing internal items that need *interpretation* cannot succeed.
2. They disagree on what conclusion to draw. Wright then goes on to work within this conception to build a concept of meaning that does not require interpretation. One does not interpret one's past or present intentions but builds them from subsequent avowals. (Although it is arguable that this is itself akin to an 'inferentialist' conception.) McDowell, by contrast, rejects the assumption that both Kripke and Wright share about mental states or items. There is no need to take this restrictive and philosophically charged attitude to the nature of mental states.

But if McDowell is right, what light can be shed on the nature of mental states? What is the alternative to representationalism?

The clue that Wittgenstein gives instead is that meaning should be explained through practice. But to try to clarify what difference invoking practice makes to an account of intentionality that can serve as an alternative to representationalism, we will now

turn to examine two contemporary philosophers: Dennett and Davidson.

The connection between Wittgenstein and Davidson and Dennett

Both Dennett and Davidson share a fundamentally third personal account of intentionality. They both set about explaining the nature of intentional mental states by examining how such states are ascribed from a third person perspective. Dennett discusses the Intentional Stance while Davidson discusses Radical Interpretation. A key issue in assessing their views, however, will be closely related to that between McDowell and Wright.

Interpretation

McDowell and Wright disagree on the kinds of things that can be present to the mind, on whether mental states can be essentially connected to worldly states of affairs or whether they can only connect once they have been so interpreted. A key issue in the exegesis of Dennett and Davidson is the nature of the input for explanation and characterization via the Intentional Stance or Radical Interpretation. Is the input bare sounds and movements, which are then *interpreted* as being about something, as having meaning or being actions with purposes? Or is it that the input must always be thought of as intentionally characterized and ineliminably norm-laden? This question will be raised first with respect to Dennett's account and then investigated more thoroughly in the context of Wittgenstein-inspired criticism of Davidson. By answering it, fresh light will be shed on the kind of practical reorientation, which stands as an alternative to reductionist naturalism.

One important preliminary point: both Dennett and especially Davidson place a great deal of weight on the notion of interpretation. We have just seen, however, that there are good reasons for rejecting a picture of mental states as free-standing internal states that need interpreting. Wittgenstein summarizes this by saying that understanding is not a matter of interpretation. Thus it may seem that Wittgenstein has already refuted Davidson's whole approach. This is an important question. But for now it is worth noting two points:

1. The Intentional Stance and Radical Interpretation do not turn on the interpretation of *internal* mental items such as signs or symbols but are a matter of making sense of the behaviour of other people. Persons and their behaviour are the object of interpretation rather than bizarre denuded internal mental objects.
2. Even a Wittgensteinian account requires some explanation of how it is possible to have access to other minds and interpretation of behaviour seems at least a plausible start.

It is a matter of live debate whether the difference between the interpretation of inner mental objects and the behaviour of people is sufficient to defuse Wittgenstein's criticisms but we will put this issue aside initially.

Reflection on the session and self-test questions

Write down your own reflections on the materials in this session drawing out any points that are particularly significant for you. Then write brief notes about the following:

1. What is the relation between rules, meanings, and mental states?
2. How different are the interpretations of Wittgenstein put forward by Kripke, Wright, and McDowell?
3. Do they target the same thing?
4. What is Kripke's sceptical argument and can it be defused?
5. What positive lessons can be learnt? What is the connection to cognitivist approaches?

Session 3 Dennett and the Intentional Stance

Dennett's carving up of the philosophy of mind

Daniel Dennett, an empirically minded philosopher, has written on three main areas within the philosophy of mind (as well as other subjects). His account of consciousness is set out in *Consciousness Explained* (1991a). *Elbow Room* (1984) sets out his views on freedom of the will. *The Intentional Stance* (1987) gathers his papers on intentionality. These are the papers that are relevant here. More recently he has also written on Darwinian evolutionary theory.

The first reading in this session is the most definitive statement of Dennett's views on the problem of accounting for intentionality in the philosophy of thought and language. In it he sets out his version of an interpretation based, non-reductive, theory of the nature of mental content. As will become clearer it is rather a minimalist theory, although not quite so minimalist as it might at first appear.

EXERCISE 4

(30 minutes)

Read the extract from:

Dennett, D. (1987). True believers: the Intentional Strategy and why it works. In *The Intentional Stance*. Cambridge, MA: MIT Press, pp. 13–35. Reprinted in Rosenthal, D. M. (1991). *The Nature of Mind*. Oxford: Oxford University Press, pp. 339–349. (Extract: pp. 340–341)

Link with Reading 25.2

- ◆ What, according to Dennett, is it to be the possessor of content-laden mental states?
- ◆ To what extent is Dennett's account of mind and meaning realist?

Realism versus instrumentalism

Dennett suggests that there is a traditional distinction between realism and instrumentalism about mental states, which can be summarized as follows:

- ◆ *The realist strategy*: mental states are literally states in the head, whose existence is independent of our description. Mental states on this view are like micro-physical entities postulated by physical science. It is thus a completely objective matter what mental state someone has.
- ◆ *The instrumentalist strategy*: beliefs are 'merely' a product of our methods of interpreting each other. Such methods do not describe pre-existing entities and are interest relative. Hence the objectivity of ascriptions is doubly questionable.

Dennett suggests, by contrast, that his account is broadly realist without construing beliefs and other mental states as internal states. He claims that this follows providing he can show that the adoption of an interpretative stance based on ascribing mental states is an objective matter. (Of course it is one thing for him to say that he is realist. It is a matter of assessment whether he is successful in this splitting the difference between the two extremes.)

Two kinds of philosophy

Dennett (1987) also distinguishes between two sorts of philosophical clarification found in the philosophy of mind. One is a form of conceptual clarification while the other relies on the development of an underlying causal theory. The distinction does not require that causality plays no part in analysis. (Some concepts essentially involve causality as a part of their analysis. To be an autograph, for example, requires a particular kind of causal history. But for most, causality need not play a part in the analysis of mental concepts.) Thus there are two distinct kinds of approach to the philosophy of mind: the conceptual analysis of intentional concepts within folk psychology and the postulation of an underlying causal mechanism to explain folk psychology. Dennett's work on the intentional stance is a contribution to the first. He suggests that he has been influenced in this by Gilbert Ryle (whose work was discussed in Chapter 22).

It is worth noting that Dennett thinks that the second strategy is a perfectly reasonable enterprise but that no elements of the brain will correspond to the elements isolated in folk psychology. Intentional system theory is a normalization of folk psychology and subpersonal theory looks at lower level workings of the brain. *But there will be no identity between elements of these two accounts.*

A key distinction

In order to clarify the distinction between these two approaches to the philosophy of mind, and to flesh out the form of realism he intends to support, Dennett borrows a distinction from the

philosopher of science and logical positivist Hans Reichenbach (1891–1953). This is between:

- ◆ *abstracta*: calculation bound ‘entities’ or logical constructs such as centres of gravity; and
- ◆ *illata*: posited theoretical entities such as electrons

The degree of Dennett’s realism can now be described using this distinction. He holds that mental states are abstracta. They play a role in a calculus of human action in the way that centres of gravity play a role in mechanics. But they are not themselves internal states, which can play a causal role like illata. By contrast, representationalists like Fodor, are, in Dennett’s phrase ‘industrial strength realists’ because they do posit internal states with causal powers.

Dennett’s ‘stance’ approach to intentionality

If mental states are abstracta, something more has to be said about the theoretical context in which they have their life. Dennett (1987) suggests that folk psychology, the network of lore that enables us to make sense of, explain and to some extent predict one another, should be clarified as a form of Intentional Stance. He clarifies this in turn, by contrast with two other stances.

Physical stance

One stance that one can take to predict and explain items or systems in the world is the physical stance. To predict the behaviour of a system determine its physical constitution, physical impingements, and physical laws. If *physicalism* is true (see Chapter 23), this stance should apply in principle to any system in the world. If all events are physical events, then the physical stance should have an unlimited field of application, again in principle. In practice, however, it is not always the best method of explaining or predicting. Suppose I wish to predict when my digital alarm clock will ring in the morning. One strategy would be to take it apart, investigate its circuits and chips, determine its voltages and currents and apply physical laws to make a prediction. In the case of an alarm clock this should work in principle because it is a physical mechanism whose behaviour is presumably fixed by (deterministic?) physical laws. In practice, however, the best way is to see what time the alarm has been set for and predict that that is the time at which it will ring. Hence the design stance.

Design stance

To predict the behaviour of a designed system ignore its physical details and predict that it will behave as it is designed to do. One adopts this stance for *pragmatic* reasons in the case of designed systems. Clearly, as not everything has been designed it has limited application. Even in the case of designed systems, it fails to apply when things break down and those breakages cannot be explained from this stance.

Intentional stance

Finally, in the case of some systems, even the design stance is ‘practically inaccessible’ and Dennett advocates instead the

Intentional Stance. One adopts this in the case of rational systems. Determine what beliefs and desires it ought to have given its position in the world and determine what behaviour would further its ends and predict that it will behave like that. (Note that these are not so niche sensitive as the design stance.) Again the initial justification is pragmatic: one adopts it for practical ease.

The kind of explanation implicit in folk psychology as explained as the Intentional Stance is:

- ◆ *justificatory*: explanations of behaviour provide rational justifications of it with respect to the beliefs and desires ascribed.
- ◆ *idealized*: its explanations do not depend on subsumption under laws but by comparison with an ideal (cf. chapter 23 for a similar claim used as an objection to functionalism).
- ◆ *abstract*: in that the states postulated need not be intervening distinguishable states of an internal causal system.
- ◆ cashed out in terms of *evolutionary* biology. Dennett offers a biological explanation of why we are the sort of creatures for whom the Intentional Stance is predictive.

Dennett’s bald key claim

Having set out this preliminary ground work, Dennett can then advance his central claim about the nature of mental states. In ‘True believers’ (1987) he sets out a bold central claim and then modifies it to take account of potential criticism. *To be a believer is to be explained by the Intentional Stance*. This view aims at the kind of relaxed view about entities which Pragmatists in the philosophy of science (such as James, Dewey, and Pierce) advocated. In Dennett’s hands the view begins something like this: there is nothing more to the question as to the existence and nature of content-laden mental states, or intentionality in general, than the question of whether the Intentional Stance applies and what its nature is. As it is described simply as a useful way of predicting some ‘systems’ in accordance with an idealized rational structure, it seems that there is no specific philosophical worry left. Note also that such an account places meaning, construed in the broad sense of Sabat and Harré firmly in the public sphere. It is to be understood as internal to a public method of interpreting one another.

Despite the simplicity and appeal of this bald approach, Dennett is, however, forced some way away from such a fully pragmatist and minimalist account.

Dennett’s bald claim modified

Further qualifications to the bald claim

The bald first statement of Dennett’s view is of very wide scope and includes animals, chess playing computers, and plants, and even lightning. Keep in mind the attraction of such a bald claim. Like pragmatism about theoretical entities in the philosophy of science, it says that all there is to the question of the real existence of beliefs is that they appear in working theories of phenomena and there is no higher metaphysical standard. There is no further

question about the realism of theoretical items beyond the fact that they appear in successful theories. But this advantage has to be diluted by responding to a number of objections that Dennett (1987) anticipates in 'True believers' (and elsewhere).

Objection 1

The crude definition is manifestly too wide because one can ascribe beliefs and desires to a lectern to predict its behaviour. Perhaps it desires to remain unnoticed and believes that by keeping still it will remain unnoticed. If the bald claim were correct then a lectern would have mental states as its staying still when not moved might be explained by ascribing to it the desire not to be noticed and the belief that by staying still it can best realize this.

Dennett's response

In response to this thought, Dennett narrows the claim as follows: To be a believer is to be predicted *with pragmatic advantage* by the Intentional Stance. We gain nothing by invoking the Intentional Stance in the lectern case that we did not have via the physical stance. The physical stance provides as quick a predictive hold.

Objection 2

Pragmatic advantage is, however, perspectival. It depends on what other explanatory resources are also available. Clever Martian neurologists and physicists might be able to predict our behaviour without deploying the Intentional Stance. So, *for them*, we have no mental states. So possession of beliefs is not an objective matter but depends upon the perspective of the interpreter.

(Note that for committed social constructionists this might be an acceptable view. They might well agree that whether I have mental states might always be a relative matter, depending on what you, for example, say about me.)

Dennett's response

Dennett, however, denies that there is no advantage even to the Martians of adopting the Intentional Stance. Without it, the Martians miss the *pattern* that is present in intentional action. Different ways of greeting share the fact that they are all greetings but lack any common physical basis. The same broad behaviour can be realized in different physical movements. This is the same point that was deployed as a criticism of type identity physicalism in chapter 23. (Compare this with what we can call 'financial transaction physicalism'. This is the claim that all monetary units are realized by physical states of some sort or other. But what physical properties do all the possible realizations of £500 have in common? Some are collections of coins, others notes, others handwritten cheques, others computer memory stores.) Thus even in the case of the Martian physicists, the Intentional Stance has pragmatic advantage. It is the only stance from which these patterns can be described. Thus there is no threat of perspectivalism and nothing wrong with Dennett's account of mental states.

Objection 3

Of course any such response raises an obvious counter response. One might argue that the Martians miss nothing by not adopting

the Intentional Stance. Their predictions as to human sounds and movements, for example, are perfectly exact even though they cannot hear meanings in those sounds or see actions in those movements. In other words one might bite on the bullet and insist on an instrumental reading of the Intentional Stance such that the intentional pattern the Martians are supposed to miss is just an *artefact* of the theory and has no theory independent existence.

Adjudicating this disagreement largely depends on what sense can be given to the idea that there really are patterns in the affairs of humans which are not merely artefacts of our view of one another. One line of inquiry would be to press the claim that there is something wrong with the explanatory apparatus implicit in the Intentional Stance. This is the claim made by eliminativists that intentional vocabulary is misleading and the generalizations of folk psychology are largely false. (See Fodor's defence against such an attack in Chapter 24.) Dennett himself goes on to defend the reality of patterns in general and thus claims that the patterns of action are simply one case of that general phenomenon and we will now turn to a further paper of his. But it is also worth briefly flagging a related debate to which we will return in the next session on Davidson.

The next session will discuss Davidson's account of 'Radical Interpretation' and a Wittgenstein-inspired criticism of it. The criticism will turn on whether in interpretation, meaning or intentionality is read into bare sounds and movements, or whether in fact that conception of access to other minds is incoherent. If it is incoherent then the alternative view is that sounds and movements have always to be described in intentional terms. The same kind of worry and response can be made in Dennett's case. The worry is that if the Intentional Stance is construed as a way of reading meaning into bare sounds and movements that do not have any intrinsic meaning, it will face a similar Wittgensteinian objection. Thus one way of seeing Dennett's defence of real patterns in human behaviour is a way of heading off that criticism.

Thus while a first way of reading the Intentional Stance is as a way of interpreting more primitive hard behavioural data by reading into them an intentional pattern, in fact Dennett is driven to claim that the patterns are there anyway, whether we detect them or not. A similar transition seems to be present in Davidson's work. While the early Davidson (and some of his contemporary commentators) suggested that the input to Radical Interpretation comprised bare sounds and movements, the later Davidson suggests that this is not possible. Once the input is denuded of its meaning, there is no way to recover it. As we will see, this point is a reiteration of Davidson's general claim that intentional notions cannot be reduced to non-intentional notions. But we will return to this issue at the end of Session 4.

The Intentional Stance *describes* 'real patterns'

Although Dennett is classed as an instrumentalist in some textbooks and collections on the philosophy of mind, that places him too far from realism. Instrumentalism is, after all, usually characterized as the view that theoretical entities are postulated *merely* in

order to save the observable phenomena and that theoretical statements are (literally construed) neither true nor false. But Dennett does not subscribe to such an antirealist view of mental states. He does not think that the pattern described from the Intentional Stance is merely in the eyes of the beholder. Dennett develops the idea of there being a real pattern in human action, which is *described* by the Intentional Stance in his paper 'Real patterns' (1991b).

Dennett starts by pointing out that we can use folk psychology to interpret the speech and actions of one another. But we can also use it successfully to predict action. Whence this power? Without some sort of pattern to events, nothing is predictable. So Dennett takes it upon himself in this paper to characterize further the kind of pattern that is on the one hand, there to be perceived from the Intentional Stance, but is on the other invisible from the Physical Stance.

Returning to the idea that mental states are abstracta (like centres of gravity, by contrast with illata like electrons), Dennett suggests that debate about whether abstract objects are real can take two separate paths:

1. metaphysically, do abstract objects in general exist?
2. scientifically, is a specific putative abstract object good in the sense of useful?

Dennett suggests, in a pragmatist tradition, that only the second is a useful and therefore good question. (It concerns usefulness after all.) Thus he suggests that the efficacy of the Intentional Stance can be attributed to the fact that the patterns it picks out are real, in the sense of being scientifically useful patterns. That is, the standard of reality to be appealed to. That is why Dennett invokes Fine's Natural Ontological Attitude, which insists on keeping such questions at the natural scientific level (see Chapter 13 for a discussion of Fine's 'NOA').

What is a pattern?

Using the analogy of a 'bar code' pattern of a series of black and white stripes partly degraded with black spots in the white bars and vice versa, Dennett (1991b) suggests that there is a real pattern 'if there is a description of the data that is more efficient than the bit map, whether or not anyone can concoct it' (p. 34). In the example at hand, it is more efficient to say that there is such and such a bar code of overall black and white stripes and then to pick out the exceptional spots than to describe each pixel individually. Thus there *really is* a pattern there to be perceived, despite the 'noise' obscuring the underlying 'signal'.

This talk of a real pattern obscures a further complication: perhaps there are several different ways of describing the fudged bar code, which differ in the way they pick out further patterns in the way apparently exceptional spots litter the wrong stripes. Dennett suggests that within this rivalry, decisions have to be made on pragmatic grounds of ease of use. This suggests a limit to the finality of questions of just *which* pattern is real. There may be no unique answer.

To give a particularly rich example of a pattern, Dennett discusses the 'Game of Life' developed by John Horton Conway,

John von Neumann Professor of Mathematics at Princeton University. The Game of Life is an example of a 'cellular automaton'. It is 'played' on a chessboard-like array of squares and is supposed to depict the life of cells on that array. Each cell is either briefly on or off. Each has eight neighbours. Three simple rules govern whether in the next 'turn' the cell is on or off. If two of its neighbours are on then it stays in its present state (whether off or on). If three are on then it turns on. For any other number it turns off.

Given these simple rules some initial shapes give rise to further stable patterns while others die away. Some patterns reproduce themselves slightly to one side and thus seem to move across the board. These are called 'gliders'. There are a number of other recognizable patterns. In principle, also, there are shapes that represent the functioning of a Turing machine or abstract computer.

This suggests different levels of description are available. One can describe the individual pixels of the Game of Life board. Or one can describe the patterns. Describing the pattern of development of patterns in the two-dimensional array of pixels using the vocabulary of stable shapes such as 'gliders' or more dramatically as describing an array of pixels as representing a Turing machine gives one much easier predictive power than relying on the piece-meal calculation of each generation of pixels. This provides an analogy for predictions of human behaviour using the Intentional Stance:

The scale of compression when one adopts the intentional stance toward the two-dimensional chess-playing computer galaxy is stupendous: it is the difference between figuring out in your head what white's most likely (best) move is versus calculating the state of a few trillion pixels through a few hundred thousand generations. But the scale of the savings is really no greater in the Life world than in our own. Predicting that someone will duck if you throw a brick at him is easy from the folk psychological stance; it is and will always be intractable if you have to trace the photons from brick to eyeball, the neurotransmitters from optic nerve to motor nerve and so forth (Dennett, 1991b, p. 42).

So is Dennett a realist then?

This leads back to the question of how realist or instrumentalist Dennett is. He himself provides a comparison with Fodor, Davidson, Rorty, and Churchland. Using the analogy of the hard lines of the bar code pattern for the 'determinacy' of ascriptions using the Intentional Stance he comments:

Fodor and others have claimed that an interior language of thought is the best explanation of the hard edges visible in 'propositional attitude psychology'. Churchland and I have offered an alternative explanation of these edges... The process that produces the data of folk psychology, we claim, is one in which the multidisciplinary complexities of the underlying processes are projected *through linguistic behaviour*, which creates an appearance of definiteness and precision, thanks to the discreteness of words.

So whereas Fodor's 'industrial strength realism' requires that the pattern detected by folk psychology is a reflection of a

pattern in the head, Dennett thinks that there may be no underlying pattern *there*. This means that no such pattern in the behaviour of neurological states can ever be detected by scanning techniques. Instead, the pattern of meaningful action, although caused by a variety of processes many of which are in the head, occurs at no deeper level than is observed in daily transactions.

The Intentional Stance and discursive psychology

So much then for Dennett's third person approach to meaning and intentionality. It provides one way of construing the claim in Sabat and Harré that meaning is constituted in conversations. But it is a weaker claim than those authors probably support. For Dennett, meaning is best understood as connected to its own special explanatory stance: the Intentional Stance. But this does not mean that meaning is merely read into otherwise intrinsically meaningless sounds and movements. The pattern of meaning is there all along and is *described not constructed* by the Intentional Stance. Nevertheless it is a necessary feature of content-laden mental states that they are accessible to others because just what we mean by intentionality is that for which the Intentional Stance is the right strategy. To refine our understanding of this view, and the role of rationality in it, we will turn to another philosopher who shares many of the assumptions of Dennett: Donald Davidson.

Reflection on the session and self-test questions

Write down your own reflections on the materials in this session drawing out any points that are particularly significant for you. Then write brief notes about the following:

1. How does thinking about the Intentional Stance clarify the nature of meaning?
2. Are mental states real objects?
3. Are mental states, according to Dennett, merely aspects of a *theory* of human behaviour or are they independent of that?
4. Does the Intentional Stance construct or describe mental states?

Session 4 Davidson and Radical Interpretation

Davidson's overall philosophical project

Donald Davidson's (1917–2003) philosophy can be roughly classified as concerning the metaphysics of mind (early work on which is largely gathered in his *Essays on Actions and Events*, 1980) and the philosophy of thought and language (early work on which is largely gathered in his *Inquiries into Truth and Interpretation*,

1984). We have already met his contribution to the former: anomalous monism and the token identity theory, in Chapter 23. This chapter will focus instead on his account of content based around an account of 'Radical interpretation' (1984 pp. 125–141).

A full understanding of Davidson's account of intentionality would require a detour into the philosophy of language proper and Tarski's theory of truth. But that is not necessary for present purposes and the discussion here will focus on what another philosopher, and influential interpreter of Davidson, Richard Rorty has described as Davidson's 'philosophy of language of the field linguist'. We will say a little more about Davidson's relation to Tarski below. The reading sets out the key ingredients of Davidson's approach to the philosophy of content. It is quite difficult on a first reading. Guides to Davidson's philosophy are listed in the Reading guide at the end of this chapter.

EXERCISE 5 (30 minutes)

Read the extracts from:

Davidson, D. (1984). Radical interpretation. In *Inquiries into Truth and Interpretation*. Oxford: Oxford University Press, pp 125–141. (Extracts: 125–126, 127–128, 136–137.)

Link with Reading 25.3

- ◆ What is the connection between ascriptions of beliefs and meanings?
- ◆ What is the evidence for Radical Interpretation?
- ◆ What is the role of rationality?

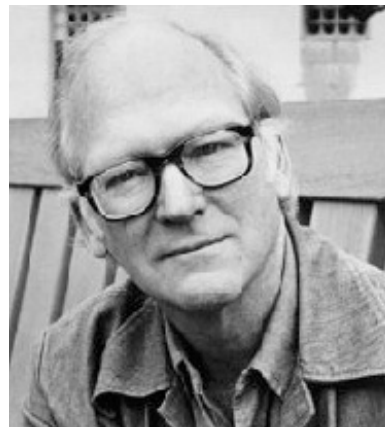


Fig 25.2 Donald Davidson

The nature of Radical Interpretation

Davidson's approach to content is based on the thought experiment of Radical Interpretation. In order to clarify what we understand when we understand our home language, Davidson considers the conditions of possibility of the Radical Interpretation of a foreign language. Radical interpretation is supposed to be interpretation from scratch. It is a philosophical abstraction from the kind of interpretation undertaken by a field linguist having first contact with an alien tribe. Such interpretation—it is assumed—cannot appeal to bilingual speakers or dictionaries. It precedes those resources. Furthermore, according to Davidson, it cannot make substantial use of the content of the mental states of speakers. Whatever the connection between mental content and linguistic meaning, Radical Interpretation must earn access to, and cannot simply assume, facts about both.

Instead, interpretation must rely only on the evidence of correlations between utterances and the circumstances which prompt them. As he says elsewhere: '[The radical interpreter] interprets sentences held true (which is not to be distinguished from attributing beliefs) according to the events and objects in the outside world that cause the sentence to be held true.' (Davidson, 'A coherence theory of truth and knowledge', p. 317).

Davidson and Quine

The idea of thinking about thought from the perspective of an anthropologist, which is central to Davidson's account of Radical Interpretation is a development from the American pragmatist philosopher W.V.O. Quine's (1908–2000) account of radical *translation* which Quine discussed in an influential book *Word and Object* (1960). There are, predictably, similarities between the accounts. But, without assuming knowledge of Quine, two key differences help shed light on Davidson's project. One is that Quine hoped that the thoughts ascribed by the imaginary anthropologist could be correlated by him or her to bodily stimuli construed as what things occur at the boundaries of the subject's body or proximal stimuli. The other is that Quine thought that the anthropologist need presuppose no mentality on the part of the subject and earn the right to ascribe mentality on the basis of mere descriptions of the subjects bodily reactions. Davidson rejects both of these behaviouristic and reductionistic assumptions (although he calls them 'details!'):

The crucial point on which I am with Quine might be put: all the evidence for or against a theory of truth (interpretation, translation) comes in the form of facts about what events or situations in the world cause, or would cause, speakers to assent to, or dissent from, each sentence in the speaker's repertoire. We probably differ on some details. Quine describes the events or situations in terms of patterns of stimulation, while I prefer a description in terms more like those of the sentence being studied; Quine would give more weight to a grading of sentences in terms of observability than I would; and where he likes assent and dissent because they suggest a behaviouristic test, I despair of behaviourism and accept frankly intensional attitudes toward sentences, such as holding true

Davidson (1984, p. 230).

Davidson takes the evidence available to Radical Interpretation to be worldly facts and events in the environment of speakers together with the occasion of their utterances. The role that evidence plays is important and we will return to this issue having sketched in the underlying purpose of Radical Interpretation.

Davidson's methodology

Davidson's methodological claim for the philosophy of content is that one can clarify the nature of both linguistic meaning and mental content more generally by examining how it is determined in Radical Interpretation. 'What a fully informed interpreter could learn about what a speaker means is all there is to learn; the same goes for what the speaker believes.' ('A coherence theory' p. 315). Because it is intended to serve this philosophical purpose, Davidson concentrates on clear instances of Radical Interpretation—interpretation by field linguists—rather than the 'interpretation' that, he claims, takes place in daily life: 'All understanding of the speech of another involves Radical Interpretation. But it will help keep assumptions from going unnoticed to focus on cases where interpretation is most clearly called for: interpretation in one idiom of talk in another.' (*Inquiries*, pp. 125–126).

Nevertheless, Davidson also thinks that everyday understanding of language involves Radical Interpretation. That claim puts some strain on the initial characterization of Radical Interpretation as interpretation from scratch because it undermines the contrast that such a description presupposes. If everyday 'interpretation' is also really from scratch, what example could there be of interpretation which was not? But while Davidson makes this claim in part to defend his radical thesis that communal language plays no explanatory role in human understanding, it can also be seen as a reminder of the purpose of considering Radical Interpretation. That is to shed light on what is understood when we understand speech and action generally. (We will return to the question of whether everyday understanding can be called interpretation at all.)

Seen in this light, Davidson's account of Radical Interpretation serves as an example of reconstructive epistemology. It does not matter that our everyday understanding of other speakers does not proceed using the tools that Davidson describes. One might argue that everyday understanding works on the implicit and tacit assumption that others speak the same language as oneself. But Radical Interpretation does not aim at phenomenological accuracy. Similarly, it would not matter if real field linguists made use of interpretative heuristics less minimal than those Davidson describes. An example of that might be the assumption that any newly encountered human language has a good chance of being related to some previously encountered language. Such a principle would be useful if it turned out that all human languages sprang from a common source. As Radical Interpretation is really a piece of reconstructive epistemology, it concerns the ultimate *justification* of ascriptions of content whatever the actual process of reasoning that gives rise to them. It concerns the evidence that could be used to justify both the possible heuristic suggested above and also our everyday methods of understanding. Radical interpretation is supposed to explain what the assumption that other speakers

speak the same language amounts to. (According to Davidson, one of its consequences is that such talk of shared languages is of no philosophical significance.) It is precisely because it plays a clarificatory—via a justificatory—role that Radical Interpretation is characterized in the austere terms that it is.

The connection between meaning and mental states

Davidson thinks that, ultimately, facts about mental content have to be determined in the same way as facts about linguistic meaning. The meanings of words and the contents of beliefs are interdependent. This presents a principled difficulty for Radical Interpretation:

A speaker who holds a sentence to be true on an occasion does so in part because of what he means, or would mean, by an utterance of that sentence, and in part because of what he believes. If all we have to go on is the fact of honest utterance, we cannot infer the belief without knowing the meaning, and have no chance of inferring the meaning without the belief. (*Inquiries* p. 142)

Thus the interpreter faces the task of unravelling two sets of unknowns—facts about meaning and facts about beliefs—with only one sort of evidence: linguistic actions which depend on both meaning and belief. How can the interpreter—to change the metaphor—break into this interdependent set of facts?

Davidson's twofold solution

Davidson's solution has two ingredients. First, he takes the evidential basis of Radical Interpretation to be the prompted assent of a speaker, which he characterizes as 'the causal relation between assenting to a sentence and the cause of such assent.' The reason for this is that it is possible to know that a speaker assents to a sentence without knowing what the sentence means and thus what belief is expressed by it (or vice versa). Characterizing a speaker as holding a particular sentence true is an intentional interpretation of what is going on—the speaker is described by relation to a propositional content—but it does not *presuppose* a semantic analysis of the sentence. That will be derived later.

The Principle of Charity

The second step is to restrain the degrees of freedom of possible beliefs in order to interpret linguistic meaning. The interpreter must impose his or her own standards of truth and coherence on ascriptions of beliefs and meanings. There must be a presumption that any utterance or belief held true really is true. Further, in a significant range of cases, the interpreter must assume that the object of an utterance, and the belief the utterance expresses, is the cause of the utterance and belief. (As Davidson remarks in a passage quoted above, the relevant cause is a worldly state of affairs rather than, as Quine (1960) suggests, proximal stimulation at the boundary of the body.) This complex of related assumptions governing the rationality imputed—generally briskly labelled the Principle of Charity—enables interpretation to get off the ground. If utterances are assumed by the interpreter to be generally true and to concern the worldly states of affairs that prompt them, then they can be correlated with those observed states of affairs. Their meaning can thus be determined. Given an overall interpretation, exceptional false beliefs can then be identified.

Holism

These a priori constraints on interpretation operate in a general manner but allow exceptions. Thus even the basic datum that a speaker holds a particular utterance true can be revised in the light of the subsequent interpretation of their other beliefs and meanings. The epistemology of interpretation is fallible and holistic. So the appeal to evidence should not be regarded as a foundational or reductive account of meaning.

(If language mastery is also holistic, such that one understands cluster of concepts, what happens if it breaks down in the case of Alzheimer's? We should expect great difficulty in understanding the speech of such sufferers. See Schwartz, M. (1990).)

The central role of rationality

But even though the constraints that Davidson points towards allow exceptions, they do suggest the following general constraint on possession of a mind. Only creatures whose behaviour and speech responses fit a general rational pattern towards others and towards the inanimate world can have minds. Minds are essentially generally rational. On the assumption that all the facts about the meanings of utterances and the contents of mental states (or about intentionality) are available to a third person radical interpreter, and on the assumption that Radical Interpretation is only possible by presupposing a largely rational pattern of behaviour, then possession of intentionality or mindedness in general presupposes an underlying rationality. The first assumption is justified by the claim that it is fundamentally our predicament if asked to justify our claims to know other people's minds, including those who taught us language. The second is justified by consideration of the constraints on the success of Radical Interpretation. Davidson's conclusion is that rationality plays a *constitutive* role in what we understand by having a mind. We could not understand a creature as having a mind if their behaviour did not display such rationality. This claim is sometimes called the Constitutive Principle of Rationality.

An overview

Davidson's basic strategy can now be summarized as follows. On the assumption that Radical Interpretation has access to all the facts about content, content can be explicated by examining the conditions of possibility of Radical Interpretation. Thus Davidson assumes that content can be captured by a third person perspective and that it can be fully analysed through its connection to the action of agents in the world. In the weakest sense of the term, Davidson can be seen, in his philosophy of content at least, as promoting a form of philosophical behaviourism providing that this is not construed in its Quinean and reductive sense. Meaning is explicated through its role in human behaviour. (In fact, he adds to this picture of content a token identity theory in order to explain the *causal* role of content to which was discussed in Chapter 23.) This is why Rorty describes this basic approach as the 'philosophy of language of the field linguist'.

Box 25.1 Davidson's formal theory of meaning

A fuller understanding of Davidson would also require a detour into his more formal work. This box provides a brief optional sketch of how Davidson construes a theory of meaning. The Reading guide is listed at the end of the chapter.

Davidson suggests that the output of the process of Radical Interpretation can be regimented in a formal theory of meaning. He assumes that the theory can be extensional and employ merely the first order logic employed in Tarski's account of truth.

Although it plays a central role in his philosophy of language, Davidson fails to explain the purpose of the theory of meaning. What is it that such a theory explains? He makes two comments on the subject. One is that knowledge of such a theory would suffice for understanding (*Inquiries*, p. 125). The other is that it is a necessary condition for languages to be learnable that a constructive or compositional account of the language could be given (p. 3). But even taken together these do not explain how provision of a theory of meaning helps the philosophical enterprise of clarifying linguistic and mental content.

He is more explicit in his reasons why such a theory should be extensional. Theories of meaning of the form: *s* means *m*—where *m* refers to a meaning of a word or sentence—have proved to be of little use in showing how the meaning of parts of a sentence structurally determine the meaning of the whole. Things can be improved by modifying the theory's structure to be: *s* means that *p*, where *p* stands for a sentence. But this still leaves the problem that 'wrestling with the logic of the apparently non-extensional "means that" we will encounter problems as hard as, or perhaps identical with, the problems our theory is out to solve.' (p. 22). The solution is to realize that what matters for such a theory is not the nature of the connection between *s* and *p* but that the right *s* and *p* are connected:

The theory will have done its work if it provides, for every sentence *s* in the language under study, a matching sentence (to replace '*p*') that, in some way yet to be made clear, 'gives the meaning' of *s*. One obvious candidate for matching sentence is just *s* itself, if the object language is contained in the meta-language; otherwise a translation of *s* in the meta-language. As a final bold step, let us try treating the position occupied by '*p*' extensionally: to implement this, sweep away the obscure means that, provide the sentence that replaces '*p*' with a proper sentential connective, and supply the description that replaces '*s*' with its own predicate. The plausible result is

(T) *s* is T if and only if *p*. (p. 23)

Further reflection suggests that, if this is to serve as an interpretation, the appropriate predicate for T is truth. We want the sentence *s* to be true if and only if *p*.

The proposed theoretical schema has the further advantage (and motivation) that it dovetails with Tarski's account of

truth. Tarski's account is pressed into service to show how the meanings of sentences are constructed from the meanings of words (which are themselves abstracted from the meanings of sentences). Davidson's use of Tarski inverts its normal explanatory priority. Tarski assumes that the notion of translation can be presupposed in the task of giving an extensional definition of truth in a language. By contrast, Davidson suggests that truth is a suitably primitive, transparent, and unitary notion to shed light on meaning. With this change of emphasis, Davidson can then borrow Tarski's technical machinery to articulate the structure of a given language.

Without going into its details it is worth noting one result of this strategy. Davidson replaces the intensional connective 'means that' with the extensional form *s* is true if and only if *p*. Clearly, however, the fact that the truth values of the left- and right-hand side of this conditional agree does not in itself ensure that the right-hand side provides an interpretation of the sentence mentioned on the left. In Tarski's use of the T schema, it can simply be assumed or stipulated that the right-hand side provides an interpretation by being the same sentence as, or a translation of, the sentence mentioned on the left. But Davidson has to earn the right to that claim. His suggestion is that instances of the T schema should not be thought of as interpretative in themselves (p. 61). Rather, it is the fact that each instance can be derived from an overall theory for the language, which also allows the derivation of many other instances of the T schema with the right matching of truth values, which is interpretative.

Given this regimentation, meaning is fundamentally holistic. As instances of the T schema are not interpretative in isolation, it makes no sense to ascribe meaning to elements of language in isolation from the rest. Only in the context of a language does a sentence (and therefore a word) have meaning. It is this, rather than the holistic epistemology of meaning ascription, which is the fundamental source of holism in Davidson.

Aside from the apparent benefits of escaping the intensionality of 'means that' and of the ability to make use of Tarski's formal machinery, Davidson's proposed structure for theories of meaning has another advantage. The formal machinery allows the derivation of a set of instances of the T schema. This seems to make it particularly apt for formalizing the output of Radical Interpretation because, as summarized above, that begins by assuming that uninterpreted utterances are held true. Thus it seems that this formal theory of meaning encapsulates the close relation between truth and meaning emphasized in Radical Interpretation.

It is worth noting here, however, that although Radical Interpretation and the formal theory of meaning sit fortuitously together, they are independent. Even if Radical Interpretation did not rely on the basic evidence of assertions but on imperatives instead, for example, its output might still be formalized using a theory of meaning based on Tarski. Reciprocally, Davidson's account of Radical Interpretation

might be used to explicate meaning in general—its connection to action in the world, the connection between meaning and belief—without adopting the formal theory of meaning as a representation of language. The latter option what is pursued in this chapter.

Davidson on content-laden mental states

Having sketched out Davidson's account of the connection between determining meanings and determining beliefs via Radical Interpretation, we can now turn to a discussion of the clearest account that he gives of what ascribing mental states amounts to (Davidson, 1991).

Davidson's attack on 'inner objects' or mental representations

In 'What is present to the mind?' Davidson argues against the coherence of any picture of the mind which includes internal mental objects. The picture Davidson criticizes is this. To have a propositional attitude is to have an object, a propositional object, before, or present to, the mind. These objects have two roles: 'They *identify* a thought by fixing its content; and they *constitute* an essential aspect of the psychology of the thought by being grasped or otherwise known by the person with the thought.' (Davidson, 1991, p. 198).

Davidson argues that these two roles cannot be reconciled. The problem is that we take it for granted that we have authority over the content of our own mental states. But if their content is fixed by an object which is known to the thinker, then to know the content of one's own thoughts requires that one knows which object is before the mind: 'The trouble is that ignorance of even one property of an object can, under appropriate circumstances, count as not knowing which object it is.' (p. 198).

It is this difficulty that leads to the philosophical postulation of special objects, such as Fregean senses, which must be what they seem and seem what they are. But, as Davidson points out, there simply are no such objects. Thus: 'If the mind can think only by getting into the right relation to some object which it can for certain distinguish from all others, then thought is impossible. If a mind can know what it thinks only by flawlessly identifying the objects before it, then we must very often not know what we think.' (p. 201)

This argument clearly differs from Wittgenstein's or Kripke's argument against inner mental objects. Wittgenstein's argument turns not on problems with the identification of such inner objects but on the impossibility of them serving their supposed function in constituting thoughts. As Kripke makes clear, no object before the mind could have the normative connections that content-laden mental states have to their fulfilment conditions. But despite this difference in argument, the end result is the same. Thinking a thought is not a matter of having an internal object before the mind's eye. This convergence of critical views is more than just a matter of interest. The obvious question that follows from the negative result is: What then is it to have a content-laden mental state? Davidson provides a clear general account.

Davidson's response to the critical arguments is to accept the first role of objects and reject the second:

It does not follow, from the facts that a thinker knows what he thinks and that what he thinks can be fixed by relating him to a certain object, that the thinker is acquainted with, or indeed knows anything at all about the object. It does not even follow that the thinker knows about any *object* at all. Someone who attributes a thought to another must... relate that other to some object, and so the attributer must, of course, identify an appropriate object, either by pointing to it or describing it. But there is no reason why the attributer must stand in any special relation to the identifying object; all he has to do is refer to it in the way he refers to anything else. We specify the subjective state of the thinker by relating him to an object, but there is no reason to say that this object itself has a subjective status, that it is 'known' by the thinker, or is 'before the mind' of the thinker. (p. 203)

The analogy with weights

He suggests that the ascription of propositional attitudes to people functions like the ascription of weights to objects. Objects stand in various relations of the form: weighing more than, weighing less than, weighing twice as much as. For simplicity, these relations and ratios can be represented by the use of a standard. This enables weights to be ascribed to objects directly using numbers. Thus one can say of an object that it weighs 5 kg. But this does not require the addition of *kilograms* into our ontology in addition to weighty objects. On this picture, numbers are in no sense *intrinsic* to the objects that have weight or *part* of them:

What are basic are certain *relations* among objects: we conveniently keep track of these relations by assigning numbers to the objects... In thinking and talking of the weights of physical objects we do not need to suppose there are such things as weights for objects to have. Similarly in thinking and talking about the beliefs of people we needn't suppose there are such entities as beliefs. (p. 205)

The last sentence might be taken to imply some form of eliminativism. But it is clear from the context that that is not the position that Davidson supports. Instead he offers a picture that clarifies what should replace mental representations or internal objects in the metaphysics of thought. No objects come before the mind's eye. Nor are there internal states that encode propositional attitudes. But this implies neither eliminativism nor any crude behaviourism in which mental states can be identified one-to-one with dispositions to act. To be in a mental state with a certain content is for one's behaviour to be explicable from a third person perspective using a system of propositional attitudes. (This is a necessary condition. To rule out things that do not need to be so described but which could be—such as planets—Davidson would have to add some further condition. One such further condition might be that using the system of propositional attitudes must have pragmatic advantage over a merely physical description (cf. Dennett 1987, p. 23). However, the formal project of formally specifying necessary and sufficient conditions is not Davidson's purpose.)

Davidson and Wittgenstein

The general picture that Davidson suggests resembles a Wittgensteinian account in which behaviour is explained by reference to a system of content-laden states governed by normative and rational relations. It can also be coupled with Wittgenstein's claim that one learns new behaviour when one learns a language. One learns behavioural repertoires that essentially turn on one's linguistic abilities. And one also learns to describe oneself in the language of propositional attitudes. Thus there is no prospect of reducing content-laden mental states to behavioural dispositions that could be described without the resources of the language of propositional attitudes. But all that is essentially involved in having content-laden mental states is the possession of complex practical abilities and behaviour.

It is worth thinking back to the account of mental states discussed in Chapter 23 on the metaphysics of mind. There Davidson's account of anomalous monism was described as centring on the claim that mental states are identical with (i.e. they just are) physical events. Davidson claims that the account described above is also consistent with that identity theory. But given the criticisms of the identity claim, we can see this account of mental states as a possible alternative. Mental states are essentially relational states ascribed to people from a mundane third person perspective to make sense of their behaviour.

An objection to Davidson's account?

A phenomenological objection to talk of interpretation

We can now return to the objection mentioned earlier to the central notion of interpretation in Davidson, and which is also implicit in Dennett. This objection runs as follows: both Davidson and Dennett overintellectualize the 'process' of making sense of one another and thus get the phenomenology wrong. The experience of hearing meaning in someone's utterance does not feel like interpretation.

This is a criticism raised by Steven Mulhall (1990), a Wittgensteinian philosopher, on the basis of work both by Wittgenstein and Heidegger. Although it is an important criticism of approaches to meaning based on interpretation, and thus is a constraint on any form of 'discursive psychology', it will be argued below that it is not decisive.

EXERCISE 6

(15 minutes)

Read: the extract from Mulhall, S. (1990). *On Being in the World* London: Routledge. (Extract: pp. 99–106)

Link with Reading 25.4

Davidson deploys the idea of *interpretation* to shed light on the nature of meaning and understanding.

- ◆ To what extent do we interpret one another?
- ◆ How like Radical Interpretation is everyday understanding?
- ◆ Do the differences matter to Davidson's project?

The background to Mulhall's criticism is ...

In *On Being in the World* (1990) Stephen Mulhall develops a phenomenological objection to Davidson on the basis of his (Mulhall's) reading of Wittgenstein.

The background to Mulhall's criticism is Wittgenstein's (1953) discussion of seeing aspects in part II section xi of the *Philosophical Investigations*. According to Mulhall, Wittgenstein's discussion of seeing aspects and aspect perception in the second half of the *Investigations* is of general importance. It attempts to characterize the immediacy with which we experience the significance of pictures, themes, words, actions, and the world more generally. The point of the discussion of cases of *changes* in aspect, such as Gestalt switches, is to illustrate the general nature of *continuous* aspect perception. The latter characterizes our normal immediate response to words and to the world. Forging a link with the Heideggerian notion of the ready-to-hand, Mulhall (1990) suggests that our experiences of the world are usually immediately charged with significance. They do not have to be interpreted.

... Wittgenstein's account of secondary sense

Wittgenstein describes this kind of immediate understanding of the meaning of a word in isolation as a form of understanding. But while this is not a metaphorical use of the word 'understanding' it is nevertheless a *secondary* use (Wittgenstein, 1953, p. 216). A secondary use is one which we find natural given the primary use, but which is discontinuous with, and could not be used to teach, the primary use. Nor is it metaphorical. An example is the use of 'fat' in the thought that Wednesday is fat. Clearly Wednesday cannot in any ordinary sense be compared with other fat or thin things. And it would be optimistic to attempt to teach the meaning of fat by giving Wednesday as an example. Nevertheless, many language users give spontaneous expression to the thought that Wednesday is a fat day.

Thus, although we may wish to say that a word or action can be immediately experienced as bearing a meaning in isolation, this does not contradict Wittgenstein's general connection of meaning with an extended practice or technique. The concept of meaning is used in its primary sense in the latter defining context and only in a secondary sense in the former. This distinction is important because Wittgenstein (1953) claims that, although as a matter of contingent fact it is not true of us, it would make sense to ascribe to someone understanding in the primary sense unaccompanied by the secondary aspect. He calls such a person 'aspect blind' (p. 213).

Mulhall (1990) argues that Davidson's use of Radical Interpretation to explicate meaning must be fundamentally mistaken because it presupposes that language users are all aspect blind:

[I]t is important to note that the metaphysics of the given—revealed as it is by Davidson's emphasis upon the concept of 'interpretation'—exemplifies to perfection the stance of the interlocutor in Section xi of the *Philosophical Investigations*. Incapable of finding a home for the notion of continuous aspect

perception in his framework of thought, Davidson describes the everyday phenomenon of perceiving words and other human beings as if aspect-blindness were the normal human state. His emphasis on processes of theorizing as necessary in order to organize bare sounds and movements into words and actions ... commits him implicitly to a general notion of visual perception as divided into what is really seen and what is interpreted, ie as divisible in precisely the way Wittgenstein rejects. (p. 106)

A key idea here is that no account of understanding one another that turns on *interpretation* can be right because interpreting one another is an exceptional activity rather than the norm. Normally our understanding is instantaneous and requires no such activity. To think instead that understanding is always a matter of interpretation is to subscribe to what Mulhall labels an empiricist view in which one experiences bare sense data (whether visual or auditory) and then interprets it.

If Mulhall is right, this is a serious criticism. He is charging Davidson with commitment to the same sort of picture that was criticized in chapter 12 on the theory dependence of data. (Recall McDowell called just such a picture of the 'Myth of the Given' a general form of empiricism evident outside the philosophy of science.)

There is also a good Wittgensteinian criticism of any such picture. Recall the argument that Kripke (1982) used to show that if one tries to derive meanings from meaning-free states in the head that have to be subsequently interpreted, then scepticism about meaning is the result. A similar argument can be used on the idea that bare behaviour has to be interpreted as carrying meaning. If this were so, there would be no sense of interpreting correctly. Meanings really would just be in the eyes of the beholder (as Sabat and Harré, 1994, seem to suggest). But in fact there is good reason to think that this is an uncharitable interpretation of Davidson.

A defence of Davidson

Mulhall's argument is not a decisive criticism for three reasons:

1. It presupposes an uncharitable—if common—interpretation of Davidson that is not obligatory. While the early Davidson does indeed suggest that the evidence for Radical Interpretation should be described in neutral terms, the later Davidson explicitly criticizes the picture of any evidence for a belief 'whose character can be wholly specified without reference to what it is evidence for' (Davidson 'The myth of the subjective' p. 162). The division of perception into what is seen and what is interpreted is rejected as a (the third) dogma of empiricism.
2. Furthermore, those comments, which can be understood as expressions of Davidson's token identity theory, can be reconciled with Mulhall's criticism. The fact that meaningful utterances or actions—or even, according to Davidson, content-laden mental states—are identical with physical events or states does not imply that they are *experienced* as mere physical events and only subsequently interpreted.

3. The connection between the primary and secondary sense of meaning is far from clear. The fact that we experience the 'meaning' of words in isolation—in the secondary sense of 'meaning'—is a contingent feature of the phenomenology of meaning. Consequently, Davidson could simply reply that his account of Radical Interpretation is meant to capture only meaning proper, meaning in the primary sense. The phenomenology is a further matter. This fits with our interpretation of Radical Interpretation as merely *reconstructive* epistemology.

This third point may require a little more explanation. One of the problems of Mulhall's argument is that it is not clear what the presence or absence of the phenomenology amounts to. This is because it is difficult to describe what someone who is aspect-blind lacks without impinging on the primary sense of meaning. According to Wittgenstein, such a person cannot see aspects change, cannot see a cube 'as a cube', but nevertheless can recognize a cube (Wittgenstein *Philosophical Investigations*, pp. 213–214). Likewise she cannot experience a word as bearing a meaning in isolation but can nevertheless learn its technique of use because blindness to the secondary sense of meaning is not blindness to the first. But the moral of this separation is clear. Whatever the secondary experience consists in, as it is possible to understand and use a word without it, it is not a part of content proper. Thus it is not essential to the philosophy of content.

So even if an account of meaning or content emphasizes the role of interpretation, this need not be a fatal objection providing both that what is 'interpreted' is always conceived as meaningful—so meaning isn't merely read into it—and that talk of interpretation is not regarded as a phenomenologically accurate account of our experience. It can still be useful to talk of interpretation because it sheds light on the constraints that govern the way we make sense of each other including, for example, the central role of rationality at work here. This helps shed light on what could be meant by discursive psychology. But before returning to the first, clinical, reading, there is one other useful philosophical approach to consider.

Reflection on the session and self-test questions

Write down your own reflections on the materials in this session drawing out any points that are particularly significant for you. Then write brief notes about the following:

1. What thought experiment lies at the heart of Davidson's account of intentionality?
2. What does this presuppose about the accessibility of meaning? What key terms does Davidson deploy?
3. What effect does that have on Davidson's claims about mental states and meaning?
4. Does Davidson's account aim at phenomenological accuracy?
5. What is Davidson's attitude to inner mental representations?

Session 5 Singular thought and the division between mind and world

Another challenge to representationalism and cognitivism

This final philosophical session will examine a different kind of objection to representationalist theory of mind. This comes neither from Wittgensteinian argument about the normativity of thought (and thus that no world-independent inner state can sustain the normativity of thought), nor from advocacy of the mundane third person perspective on meaning taken by both Dennett and Davidson. Instead it is drawn largely from attention to a particular kind of thought that we all sometimes have: thoughts about objects identified merely by our direct perception of them. These are generally called 'singular thoughts'. (An example is the sort of thought you have about this page in front of you if you are thinking 'That page is difficult'.)

As the discussion in this session argues, the very possibility of such thoughts shows that the idea of the connection between mind and world underlying representationalism is mistaken.

Neo-Fregean theories of thought

The first reading in this session (linked with Exercise 7) is another from the work of the American-based English philosopher John McDowell. As well as being a commentator on the work of Wittgenstein, McDowell is also a key figure in an approach to the philosophy of content called *neo-Fregean* philosophy of thought and language. It is Fregean in that, as an approach to the philosophy of thought, its central concept is that of *sense*, which was introduced by Frege to stand in contrast to that of reference.

Sense and reference

The distinction between sense and reference was discussed as part of Kripke's theory of names in chapter 23. It was introduced by Frege as part of an account of what is understood by a speaker when they understand a name such as, in Greek, 'Hesperus', or the 'Evening Star' as it is translated. Because a speaker may know when to apply 'Hesperus' and also 'Phosphorus' (the 'Morning Star') but may not know that they stand for one and the same planet (Venus, in fact), Frege concluded that what is understood when a speaker understands a name is not simply what it refers to. (If it were so, then the Ancient Greeks would counter-intuitively have understood the same thing by both 'Hesperus' and 'Phosphorus'.) Instead, he suggested, one knows its sense. As a way of fleshing out this term he suggested that the sense stands to a referent as the 'mode of presentation' construed as a viewpoint stands to what is so presented. And he gave, as an analogy, the idea of different views of the same object from different points or through different portholes.

In the example just given it is clear that the senses can be thought of as descriptive (the morning or evening star). The sense determines the object by containing a description.

McDowell following the work of the late Gareth Evans, has argued that in fact senses do not have to be construed as descriptions which specify worldly objects as we will shortly see.

Health warning

The next reading (linked with Exercise 7) is, perhaps, the most difficult piece of philosophy so far encountered in this chapter. There is much in the paper from which it is drawn that is not directly relevant to this chapter but the key message has important consequences for thinking about thought. It begins with some comments on Russell's theory of descriptions, which we will now summarize. Read these comments again after reading the article.

EXERCISE 7 (60 minutes)

Read the extract from section 8 of:

McDowell, J. (1986). Singular thought and the extent of inner space. In *Subject, Thought and Context* (ed. P. Pettit and J. McDowell). Oxford: Oxford University Press, pp. 137–168. (Extract: pp. 158–161)

Link with Reading 25.5

- ♦ Try to work out what McDowell's objections are to a Cartesian picture of mind whether Descartes' or a modern materialist variant.

The Russellian background

A key background to the paper from which the extract is taken is discussed in the first four sections. Sir Bertrand Russell distinguished between what he called *logically proper names* and merely apparent names that have quite a different semantic structure. Logically proper names feature in *singular propositions*, which for simplicity we can think of as *singular thoughts*. A logically proper name is purely referential. It takes its meaning—what it contributes to a thought in which it features—simply by going proxy for the thing it named. As a result, if such a name fails to refer—for whatever reason—the singular thought it goes to make up would not be a proper thought at all. There would be something missing. It would be a kind of nonsense.

Russell argued that there are, however, very few logically proper names. Indeed he seems to have thought that only 'this', 'that' (when referring to sense data) and perhaps 'I' were such. A consequence of this (and probably its motivation) was that the sort of reference failure just envisaged was not possible (because one could not be so mistaken about the presence of sense data or oneself). So everything else that we would normally call a name, such as 'Moses', had to be construed differently. Russell deployed a logical analysis called the Theory of Descriptions to explain how such 'names' functioned.

The theory of descriptions

This logical 'theory' was first deployed by Russell to account for non-referring descriptive *phrases*. The most famous example of

these is the sentence 'the present king of France is bald'. If an utterance of this sentence were successfully to have meaning it seems at first that it would have to be successful in referring to an individual—the king of France—and asserting of him that he is bald. But this first thought faces a problem. As there is no present king of France there would be something wrong with an utterance of the sentence now. It seems it would lack a clear meaning. To avoid this consequence, Russell analysed the sentence as making instead a conjunction of three claims:

1. There is one king of France.
2. There is no more than one King of France.
3. That thing is bald.

This conjunction of sentences (properly, utterances) is not meaningless. It is simply false because one of its conjuncts is false.

Russell suggested that apparent names such as 'Moses', which were not logically proper names, really stood for some such descriptive content as 'the leader of the Israelites'. Thus sentences about Moses could be analysed in the same way as those about the King of France.

Names, reference failure and two kinds of knowledge

McDowell (1986) points out that, although Russell realized the possibility of there being sentences or thoughts that were vulnerable to a radical form of failure—just in the case that the thing to which they refer does not exist—he deployed the Theory of Descriptions to make sure that this would never, as a matter of fact, happen. As the only sentences that contained genuine names referred to sense data, one would never be in error about them.

This distinction in the semantics of real and apparent names—in other words what they contribute to sentences containing them—was reinforced by an underlying epistemological distinction that Russell advocated. Russell suggested that thought made contact with everyday objects in two ways: either by specifying a description that the object satisfied or through direct contact or acquaintance. Hence the distinction between knowledge by description and knowledge by acquaintance. Russell supposed that one could only be directly acquainted with sense data and perhaps oneself. Thus the only logically proper names refer to these and all other 'names' are really disguised descriptions.

McDowell's extension of the idea of singular thoughts to everyday objects

McDowell (1986) suggests that the idea of object-dependent, or singular, thoughts can be taken from Russell and made more general. Instead of only being directly acquainted with sense data one can be directly acquainted with a range of everyday objects in direct perceptions and perceptual thoughts. The linguistic sign of such thoughts is a demonstrative expression of the form 'That cup is red!'

A typical visual experience of, say, a cat situates its object for the perceiver: in the first instance egocentrically, but, granting the perceiver a general capacity to locate himself, and the objects

he can locate egocentrically, in a non-egocentrically conceived world, we can see how the experience's placing of the cat equips the perceiver with knowledge of where in the world it is (even if the only answer he can give to the question where it is 'There'). In view of the kind of object a cat is, there is nothing epistemologically problematic in suggesting that this locating perceptual knowledge of it suffices for knowledge of which object it is (again, even if the only answer the perceiver can give to the question is 'That one'). So those visual experiences of objects that situate their objects can be made out to fit the account I suggested of the notion of acquaintance: abandoning Russell's sense-datum epistemology, we can say that such objects are immediately present to the mind...

(p. 140)

But as well as drawing the idea of direct acquaintance and singular thoughts from Russell, McDowell also draws on Fregean work on thought. (This is why he talks of singular *thoughts* rather than singular *propositions*.) Frege's outlook is characterized in the following way:

Frege's doctrine that thoughts contain senses as constituents is a way of insisting on the theoretical role of thoughts (or contents) in characterizing a rationally organized psychological structure; and Russell's insight can perfectly well be formulated within this framework, by claiming that there are Fregean thought-constituents (singular senses) that are object-dependent, generating an object-dependence in the thoughts in which they figure.

(p. 233)

The first part of this quotation reiterates McDowell's methodological commitment to a Fregean or neo-Fregean philosophy of thought or content. This is centred round the notion of *sense* that McDowell here describes as playing a role in characterizing a *rationally organized psychological structure*. Recall the example of Hesperus and Phosphorus above. The same person could rationally take a different view to sentences that said the same thing about Hesperus and about Phosphorus because they may not realize that the sentences are about the same thing. This 'intuitive criterion of difference' marks out the thoughts expressed by the two sentences as being, or having, different contents. So in charting a person's rationally organized psychological structure one would want to individuate their thoughts with a finer grain than simply the level of the objects to which they refer. One needs also to take account of the senses, or the 'modes of presentation' of referents, which make up their thoughts.

What makes McDowell's remarks novel (albeit he is following the work of his late colleague Gareth Evans in making this point) is that a broadly Fregean focus on sense and not reference can be combined with the Russellian idea of singular thoughts (i.e. thoughts that are object-dependent) providing that one can think of some senses as fixing their referents more directly than by a description. (Recall that the Hesperus–Phosphorus example turns on associating different descriptions with the different names.) An example of a singular, or object-dependent, thought is what is expressed by the phrase 'this cup' in the sentence 'This cup is blue'. Think of this as how the cup is thought about when one thinks 'this cup...'

Russell's resistance to this is a sign of his Cartesianism

McDowell suggests, however, that this extension of singular thoughts to cover everyday objects would be rejected by Russell as nonsense for reasons that have to do with the Cartesian assumptions behind Russell's underlying view of the mind.

In a fully Cartesian picture, the inner life takes place in an autonomous realm, transparent to the introspective awareness of its subject; the access of subjectivity to the rest of the world becomes correspondingly problematic, in a way that has familiar manifestations in the mainstream of post-Cartesian epistemology. If we let there be quasi-Russellian singular propositions about, say, ordinary perceptible objects among the contents of inner space, we can no longer be regarding inner space as a locus of configurations which are self-standing, not beholden to external conditions; and there is now no question of a gulf, which it might be the task of philosophy to try to bridge, between the realm of subjectivity and the world of ordinary objects.

McDowell is here making two corresponding points:

1. In a Cartesian picture of the mind, there is no possibility of being in error about one's own thoughts in the way that is possible if there can be singular thoughts about everyday objects. Such thoughts might fail if there were no appropriate object to have thoughts about, because for example, of hallucination. The Cartesian picture can be more or less defined by the way in which it buys immunity to error here. It construes the mental or inner as a realm that is independent of the world of objects. (Of course it must *align* with the outer world if thoughts are to be *true*. But to *have* thoughts does not depend on standing in a relation to the world.) McDowell describes this picture as of a realm of configurations, which are self-standing, not beholden to external conditions.
2. But *if* one is prepared to accept the possibility that thoughts might fail in the way outlined, there is a corresponding gain. It is no longer the case that thought is cut off from the world. It is no longer the case that thoughts are construed as free-standing mental representations, which then have to be re-connected to the world through causal or evolutionary theories, for example.

Cartesian scepticism as the source of this picture

Section 5 of McDowell's (1986) paper attempts to diagnose how this conception of the inner came about, linking it to Descartes' sceptical project. The details of that historical project do not matter here but it is worth briefly noting one idea that has a bearing on epistemology. McDowell suggests that, on a Cartesian picture of the mind, our experience is taken to be notionally separable from the world that brings it about. Imagine two cases: one in which someone directly perceives a dagger before them and another in which they hallucinate that there is a dagger before them.

On a Cartesian picture, there is something in common to both cases: an experience as of a dagger, one might say. In the former

case this mental state is caused by a real dagger while in the latter it is not. This *highest common factor theory* contrasts with McDowell's preferred view: the *disjunctive theory* of experience. In this, experience is construed as either directly of a dagger or merely an appearance of such contact. The difference is that on the latter theory, in veridical experience, there is direct access between a subject and the world. The subject is not cut off from the world by a 'veil of ideas.' This is an important distinction for epistemology but not directly relevant to this Part except in the way it reinforces the general worry about the connection of mind and world.

McDowell's first (explicit) objection to the Cartesian picture

The final paragraph of section 5 brings out more clearly the connection between the highest common factor view of experience and the Cartesian view of the mind. It presents an underlying objection to a Cartesian account of experience. If one thinks of experience as self-standing and independent of the world (as it is in the highest common factor theory) then:

This makes it quite unclear that the fully Cartesian picture is entitled to characterise its inner facts in content-involving terms—in terms of its seeming to one that things are thus and so—at all... there is a serious question about how it can be that experience, conceived from its own point of view, is not blank or blind, but purports to be revelatory of the world we live in.

So if one is in some sort of inner state that is common to both veridical experience and hallucination, what is it about that state which is about the world. If one can never have direct unmediated contact with the world of objects, how can one's inner states be about that world?

The same criticism applies to representationalism and cognitivism

Sections 6 and 8 of the paper apply these objection to a Cartesian picture of the mind to modern views. McDowell (1986) argues that modern pictures of the mind such as functionalism, representationalism, and cognitive science fall to the same underlying objection. While modern views are not dualist in the sense of embracing an immaterialist theory of mind, they nevertheless preserve the central idea that there is an inner realm (in this case literally spatially inner) that comprises self-standing items. As such they are open to the same charge of darkness.

A related objection cashed out in terms of sense

McDowell connects this objection to modern positions back to the central role of sense in neo-Fregean philosophy of thought and language. He points out that the idea of subjects having different thoughts with different senses standing for the same object in the world has to be explained in, say, representationalism by different internal happenings. It cannot be explained by what happens outside the subject because the same outside object may be involved in thoughts with different senses. But, McDowell

argues, nothing that goes on inside in the way that representationalism construes it will help either because there is no reason to think of these self-standing internal objects as carrying meaning, as being about the world. The inner realm in such accounts remains dark.

A summary so far

We can now stand back and think about the main point of this paper as far as this chapter and Part goes. McDowell attempts to undermine an underlying Cartesian assumption about the philosophy of mind that he thinks is present in the semantic theorizing of Russell and his more recent heirs (the latter are discussed more in later sections of the reading). This is the assumption that the mind is independent of the world. Once that assumption is in place, however, it makes thought's bearing on an outer world mysterious. Representationalists think that they can make some reply to this by devising causal mechanisms to explain how inner symbols become charged with meaning. But McDowell argues that such causal connections will still leave the inner world dark and meaningless.

It is this Cartesian assumption of world independence that stops Russell from drawing more general conclusions from his idea that some thoughts might indeed be object-dependent. But trapped within Cartesian thinking, Russell restricted such thoughts to those involving logically proper names whose scope was radically limited to pointing to sense data and thus not genuinely world-dependent thoughts after all. McDowell suggests that one should take seriously the idea that thoughts can be partly constituted through direct acquaintance with the world by construing such acquaintance as normal perception of objects (rather than of sense data). If so then the very idea of there being a gulf between an inner world of thoughts and an outer world of objects will be undermined. Singular thoughts are the thoughts that they are (such as the thought that that cup is red) partly in virtue of a perceptual ingredient that singles out the cup in front of me. They are essentially world-involving.

The second (implicit) criticism

This account of what is going on in McDowell's (1986) paper leaves one missing piece of the jigsaw. It concerns the difference between descriptive thoughts and singular thoughts. Now this distinction is found in Russell's logical work. But a rough and ready distinction can be brought out again making use of Frege's 'intuitive criterion of difference'. Two thoughts are distinct (i.e. they are, or have, different contents) just in the case that it could be rational for the same subject to take different views about their respective truth.

Now consider two assertions which I might now make using two different sentences. Think whether these assertions express distinct thoughts or contents:

1. 'The cup in room S2.54 of the Warwick Philosophy department is red.'
2. 'That cup is red.'

By Frege's intuitive criterion these do *not* express the same thought even if they say of the same object (the cup in my office) the same thing (that it is red). This is because it would be rational of me to accept the second but not the first if I did not realize that I was sitting in room S2.54 (although in fact I am).

The two thoughts expressed by these assertions are distinct in form. One is a *descriptive* thought that specifies the cup in question using a description: 'The cup in room S2.54 of the Warwick Philosophy department'. The other is a *singular* thought that specifies the cup in question directly from a perception (note that 'that!' is not a description). Singular thought theorists such as McDowell and Evans have argued that there is no way of specifying a descriptive content that would coincide with a singular thought such that it would never be rational to agree to one and reject the other. In other words they think that such disagreement would always be possible and thus that singular thoughts always have different content from descriptive contents.

Why singular thoughts cannot be accommodated within representationalism

With this additional piece of the jigsaw, we are, at last able to see why McDowell's view of the mind is incompatible with any broadly representationalist or cognitivist approach in which thoughts are identified with inner structures in the brain. If we, for the moment, disregard the criticism that self-standing configurations in the brain would not have content at all, and would remain dark, then it might be conceivable that descriptive thoughts could be modelled as blue prints realized by structures in the brain. That is, if we allow some content to internal symbols (how?!) they might underpin descriptive thoughts by coding descriptions that worldly objects could *satisfy*. Thus internal structures might come to be about worldly objects. (To repeat we still haven't said how even the descriptive elements get their content.) But as singular thoughts do not work like this, they cannot be fully linguistically coded in a description.

This suggests that there is a principled objection to the idea of explaining thoughts by invoking internal vehicles of content or mental representations. While such a programme might conceivably work for descriptive thoughts, it cannot work for singular thoughts because in their case there is nothing to be so coded. Worldly objects themselves partly constitute the content of singular thoughts and thus their content cannot be captured in internal symbols.

Two open questions

This argument is important because it forms one of the foci of much recent thinking about the philosophy of content in the UK. The neo-Fregean approach is obviously not without its critics. While exploring these would take us beyond the scope of this course, here are two lines of inquiry.

1. What is the role of real objects in constituting the content of thoughts? As we saw in Session 4, Davidson argues that while it is necessary for interpreters to relate subjects with objects in

propositional attitudes in order to make sense of their speech and action, it is not necessary for the subjects to have a 'psychological' or perhaps better a 'psychologistic' connection to them?

2. Can McDowell and the neo-Fregeans really reconcile the idea of acquaintance with the Fregean notion of sense? McDowell (1986) comments in the reading linked with Exercise 7 that this can be done providing one distinguishes between objects being constituents of thoughts (a Russellian idea) and objects figuring in thoughts (a neo-Fregean idea). But what is the distinction between these two. Why will senses not serve as just the sort of veil of ideas that the neo-Fregeans criticize in Cartesianism.

Summary of the 'philosophical' sessions on antireductionism

We can now return to examine the thrust of the more explicitly *philosophical* sessions in this chapter. Consider again the two challenges raised at the start of the chapter. One is the challenge to find a place in nature for meaning given that it is unlikely, to say the least, that 'aboutness' will feature in any final account of the world arrived at in physics. A reductionist response to this challenge is to attempt to explain how intentionality is itself the product of purely physical properties. Cognitivist neuropsychology shares just that aim. And like most reductionist philosophy, it starts with an assumption, motivated by an analogy with computers, that human information processing requires internal states or representations to carry that information. However, that raises the other challenge described again at the start of the chapter. The problem is that once one construes content-laden mental states as free-standing internal states standing in some causal relations, it becomes mysterious how they can also be about anything. (Trousers hanging in a wardrobe may be independent of the world but are not about it. How does adding in some causal relations bring 'light' to the inner world?) We saw one particular version of this challenge in the discussion of how inner states might share the normative properties of mental states.

The approach discussed in this chapter, broadly shared by both discursive psychology and by antireductionist philosophers of content aims to meet the first challenge without reducing intentional notions to non-intentional notions. It does this by reminding us how meaning plays a perfectly unmysterious role in the natural world even though that is not equated with the world as described by physical science. Intentional properties cannot be described by taking the physical stance but instead by taking the intentional stance, which answers to a different constitutive principle: rationality. So rather than looking for causal patterns within the head, antireductionists look to normative and rational patterns in human actions. It is this that is gestured at, with perhaps a little clumsiness in discursive psychology to which we will now turn again.

Reflection on the session and self-test questions

Write down your own reflections on the materials in this session drawing out any points that are particularly significant for you. Then write brief notes about the following:

1. How, according to Russell, can thoughts patch onto worldly objects? What theory does he advance for the analysis of descriptive thoughts?
2. What restrictions does Russell place on what can be known by acquaintance? How have they been relaxed by recent philosophers and why? What risk does this carry for the nature of thought?
3. What significance does Russell's account of thought have for the inner representation of thought and thus for cognitivist accounts of intentionality?

Session 6 Discursive psychology and Alzheimer's disease

This final session will return to the paper with which we began this chapter: Sabat and Harré's work on Alzheimer's disease sufferers as 'semiotic' subjects, but will also look at a more careful methodological textbook on discursive psychology written by Rom Harré and Grant Gillett, a neurosurgeon turned philosopher from New Zealand.

EXERCISE 8 (30 minutes)

Look back at Sabat and Harré (see the reading linked with Exercise 1, 'The Alzheimer's disease sufferer as a semiotic subject', 1994) at the start of this chapter and then this further extract: pp 150–152

Link with Reading 25.6

Think again about the range of claims made about meaning and mental states in the research paper.

- ◆ Are they consistent?
- ◆ What support do they receive from the discussion in the rest of this chapter?
- ◆ Could the empirical findings of the paper have been arrived at using different theoretical assumptions about meaning?
- ◆ What does the overview of discursive psychology in the chapter add?

What do discursive psychologists really say about meaning?

We are now in a much better position to assess the extent to which discursive psychology really does differ from the sort of

cognitivist or representationalist models discussed in Chapter 24. The question is whether a discursive model is an alternative model that is inconsistent with that other approach or whether it is complementary, differing only in emphasis, focusing on the broader context rather than what goes on in the head.

Closer attention to the article reveals that it says a number of different things about the underpinning of meaning. So as was picked out in the first session Sabat and Harré say in the abstract that 'meanings are jointly constituted by the participants in a conversation' and 'from a discursive point of view, psychological phenomena are not inner or hidden properties or processes of mind which discourse merely expresses'. These look at first as though they are saying the same sort of thing. But we can now see that strictly this is not the case.

If one adopts the view of Dennett, Davidson, Wittgensteinians, or neo-Fregeans, the second claim is necessarily true. This is the claim that psychological phenomena (or at least content-laden mental states, propositional attitudes) are necessarily the sort of things that can be described, explained or predicted from a mundane third person stance. In other words they are necessarily available to other members of the conversation (to put it metaphorically). However, that claim need not require that meanings are constituted in this activity of interpretation. One would need a further argument to establish that claim. Even talk of *joint* constitution threatens to undermine the idea that interpretation of others can be right or wrong. But being correct in their interpretation of Alzheimer's sufferers is surely something that Sabat and Harré aim at.

Some tensions in the characterization

On the other hand, despite such occasional radical claims about the social constitution of meaning, the authors also make use of a different sort of analogy. They say: 'It is our contention that a person suffering from Alzheimer's condition is like someone trying to cut wood with a blunt saw, or trying to play tennis with a racket with a warped frame. The basic intentions may be there, but the instrument for realising them is defective.' (p. 146).

But, as Tony Hope argued in a commentary on the paper, this does not sit well with the claim about the *social* constitution of meaning. To mean something by a word is a species of intention. One means or intends to use it in a particular way. So if this is intact, all that is left for other participants in the conversation to do is to *detect* it. And that is something for which Sabat's method of time compression seems particularly well suited. But in that case, the analyst is aiming simply to describe meanings that already exist, not partially to construct them.

So is discursive psychology consistent with cognitivism?

So far then it seems that discussion of the wider philosophical background has undermined the claims of discursive psychology. It either *is not or should not be* as radical as is sometimes claimed. So one possible response is to construe it as *consistent* with the approach described in Chapter 24. If the linguistic intentions of Alzheimer's sufferers are intact even if they are difficult to detect

and describe, perhaps the non-social facts that they consist in are those described in cognitivist terms.

But as both the criticisms of cognitivism discussed in Chapter 24 and the arguments described in this chapter suggest, that happy reconciliation cannot be true. While linguistic or other intentions need not await piecemeal construction by other participants in conversations, they are not the sorts of things that can consist in mental representations processed in a 'semantic system' in the head. They depend instead on normative patterns of speech and action in human transactions. They are essentially world-involving relational states that are necessarily describable from a third person perspective. They are not hidden internal states. *That* claim of discursive psychology has been supported by the consideration of this chapter.

A more careful statement of discursive psychology?

The brisk methodological remarks in the research paper are useful guides to some of the underlying beliefs of followers of discursive psychology. Presented without the caveats and restrictions of more formal methodological works, such hastier comments are often better guides to underlying thinking and assumptions. But in a more reflective philosophical text *The Discursive Mind*, Rom Harré (here with a different co-author: psychologist and philosopher Grant Gillett) presents a more cautious statement of the nature of discursive psychology, called here the 'second cognitive revolution'. At the centre of the more cautious view is the statement: 'In this sense, the psychological is not reducible to or replaceable by explanations in terms of physiology, physics, or any other point of view that does not reveal the structure of meanings existing in the lives of the human group to which the subject of an investigation belongs.' (p. 20).

This is a good statement of an antireductionist view of intentionality and also one that begins to gesture towards the reason why such reduction is impossible. It still leaves open the question more precisely *why* the structure of meanings cannot be mapped on to or reduced to the realm of law in the way that Fodor (and others from Chapter 24) promises. But we have begun to see in this chapter and in Chapter 24 some further reasons. Internal states cannot be *about* the external world because once such states are thought of as free-standing internal states they lose their normative connections to the world that cannot be reconstructed either by acts of interpretation or by merely causal resources. (Adding in teleology 'solves' the problem as long as it is construed as *presupposing* the same normative content that was supposed to be reduced. If not it is as powerless as a pure causal story.) And in this chapter we have begun to see an alternative account of the place of intentionality in nature: as revealed in one stance towards making sense of other people and ourselves.

But there are passages in the Harré & Gillett chapter that might seem to point beyond the general account developed in this chapter. Consider this:

Thus the experimenter or observer has to enter into a discourse with the people being studied and try to appreciate the shape of the subject's cognitive world. But at this point it no longer makes

sense to talk of observers and subjects at all. They are only co-participants in the project of making sense of the world and our experience of it. (p. 21)

In the context of the research article this might sound like a statement of the ongoing social *construction* of meaning. But it does not *quite* say that. Being co-participants in a project of making sense of the world might not involve being co-creators of the realm of sense. It might involve jointly tracking or *detecting* the patterns that shape human behaviour. On the other hand, the more modest reading of the passage greatly reduces the motive for saying that 'it no longer makes sense to talk of observers and subjects at all'. There is no reason to reject this distinction (just try in a practical context of psychiatric inquiry dropping the distinction!) if keeping it does not imply that only the analyst is responsible for meanings and the 'subject' is an unwitting victim. And it does not. *Both* the observer and the subject may be tracking shared meanings in the social realm. This more modest claim does not require the idea that the meanings are made up as both parties go along.

Three principles of discursive psychology

In fact, one of the striking things about Harre & Gillett's work is just how difficult it is to pin down precisely what claims about the nature of psychological phenomena it wants to advance.

The three principles which characterize the discursive turn in psychology are summarized:

1. Many psychological phenomena are to be interpreted as properties or features of discourse, and that discourse might be public or private. As public, it is behaviour; as private, it is thought.
2. Individual and private uses of symbolic systems, which in this view constitute thinking, are derived from interpersonal discursive processes that are the main feature of the human environment.
3. The production of psychological phenomena, such as emotions, decisions, attitudes, personality displays, and so on, in discourse depends upon the skill of the actors, their relative moral standing in the community, and the story lines that unfold. (p. 27).

This summary involves a slippery use of the term 'discourse'. By construing psychological phenomena as 'properties or features of discourse' it appears to make a surprising and possibly socially constructionist claim. As we normally think of language as a public phenomenon, construing psychological phenomena as features of language makes them public also. But it then reconstrues 'discourse' in an equally surprising way—as private, it is thought—which undermines this. Unlike our normal understanding of language, discourse can simply *be* private thought.

The same slipperiness is present in the next paragraph (on p. 27), which says that: 'discursive phenomena, for example, acts of remembering, are not manifestations of hidden subjective, psychological phenomena. They are the psychological phenomena... There is no necessary shadow world of mental

activity behind discourse in which one is working things out in private.' (p. 27).

This passage contains a claim that fits the general approach of this chapter. Mental phenomena do not take place in a hidden realm of mental representations. However, the passage also *suggests* a more radical claim that nothing stands behind the *expression in language* of say, remembering. (This is suggested by the phrase 'behind discourse'.) It *hints* at the view that there is no mental activity behind the linguistic act of announcing one's memory. But again it does not actually say this. What it does say is that 'acts of remembering... are psychological phenomena'. And providing one takes an 'act of remembering' *not* to mean *saying* that one remembers but actually *remembering*, then there is nothing contentious in what is actually said.

A way to interpret the reading

But perhaps the best way to read practical accounts of discursive psychology (and, in fact, the others listed in the Reading guide at the end of this chapter) is as intermediate promissory statements located between empirical work, sharing the same emphasis on the social, the external, and the linguistic, and more explicitly philosophical work on the underpinnings of intentionality.

As such an intermediate statement it serves to summarize a general claim that has received support throughout the last two chapters. Content-laden mental states are not states of the brain. Utterances do not inherit their meaning by standing proxy for such inner states. The space of reasons (again to look to Sellars' and McDowell's phrase) is an essentially world-involving and potentially social space. It is best understood from a mundane perspective that charts the meaningful behaviour of whole people going about their lives.

Such an approach does however encourage the very careful attention to language used in context set out by, eg, Sabat in the case of Alzheimer's sufferers (Sabat 2001). Such careful empirical study does not need the further more radical claim that meaning is constructed, rather than relied upon, in dialogue.

Conclusions

We can step back from the details and take stock. Discursive psychology emphasizes the importance of a wider social context for making sense of the speech and actions of subjects. As the reading shows, this emphasis can be useful for empirical reasons. But on the face of it, one need not adopt all the claims sometimes made of discursive psychology to pursue such a programme. If linguistic intentions are in tact whether or not they can be realized, then perhaps they can be investigated using a cognitivist form of psychology.

Note also that even cognitivist approaches to mind and meaning could agree that social factors were *causally* important. It is surely plausible on any account that social factors in the form of education are developmentally important for being able to think thoughts about, say the balance of payments or electrons. This is a causal dependence.

But there is at least a strand of thinking within discursive psychology that makes the further claim that relations to things

outside the head play a *constitutional* rather than merely a *causal* role. We have seen that if this is the claim that meanings are constituted through the interpretation of bare sounds then it is flawed. But if it is the claim that content-laden mental states are necessarily publicly accessible and world-involving then it receives support from a range of philosophical approaches to meaning. So the general result of the philosophical work surveyed over the last two chapters is that there is general a priori support for an approach to intentionality akin to a modest form of discursive psychology by contrast with the reductionist approaches of, say, cognitive neuropsychology. (This is not to say that every or even the majority of philosophers would agree with the arguments marshalled in these chapters.)

Reflection on the session and self-test questions

Write down your own reflections on the materials in this session drawing out any points that are particularly significant for you. Then write brief notes about the following:

1. What in general is the connection between the philosophical models of meaning provided by Wittgenstein, Dennett, Davidson, and McDowell and discursive approaches to psychology and psychiatry?
2. What support is given to social constructionism? Is social constructionism necessary to distinguish discursive psychology from cognitivism?

Reading guide

- ◆ The discursive approach to psychology is outlined in a number of places: Church (2004) 'Social constructionist models: making order out of disorder—on the social construction of madness' (in Radden (ed.) *The Philosophy of Psychiatry*); Edwards and Potter (1992) *Discursive Psychology*; Harré and Gillett (1994) *The Discursive Mind*; Sabat (2001) *The Experience of Alzheimer's Disease*; and Sabat and Harré (1994) 'The Alzheimer's disease sufferer as a semiotic subject'.
- ◆ It is further discussed in Gillett (1997a) 'A discursive account of multiple personality disorder', with a commentary by Braude and a response by Gillett.
- ◆ It is criticized from a distinct related perspective in Coulter (1999) 'Discourse and mind'.
- ◆ For an introduction to the philosophy of thought and language broadly consistent with discursive approach see Luntley (1999) *Contemporary Philosophy of Thought*, and Miller (1998) *Philosophy of Language*.

Wittgenstein on rules

- ◆ There is a very great deal written on the interpretation of Wittgenstein's (1953) *Philosophical Investigations*. A good place to start is McGinn's (1999) *Wittgenstein's Philosophical Investigations*, and Thornton's (1998) *Wittgenstein on Language and Thought*, which contains critical discussion of Kripke, Wright, and McDowell.
- ◆ For an interpretation of Wittgenstein that supports a radical social constructionism see Bloor's (1997) *Wittgenstein on Rules and Institutions*.
- ◆ See also Coulter (1979) *The Social Construction of Mind*.
- ◆ A good collection of essays on Wittgenstein's discussion of rules can be found in Miller and Wright (ed.) (2002) *Rule-Following and Meaning*.

Dennett and Davidson

- ◆ A useful beginning to Dennett's philosophy of mind is his *Kinds of Minds* (1996).
- ◆ Dennett's philosophy is discussed in Haugeland 'Pattern and being' and Rorty 'Holism, intrinsicity and the ambition of transcendence', both in Dahlbom (ed.) *Dennett and his Critics* (1993). A clear statement of Davidson's philosophy is Davidson's (1984b) 'Belief and the basis of meaning' in his *Inquiries into Truth and Interpretation*, pp. 141–154.
- ◆ Davidson's philosophy is introduced in Evnine (1991) *Donald Davidson*.
- ◆ There are useful critical essays in LePore and McLaughlin (ed.) (1985) *Actions and Events*.
- ◆ His formal theory of meaning is debated in Dummett (1993) 'What is a theory of meaning I and II' (in *The Seas of Language*), and McDowell (1999) 'In defence of modesty' (in *Meaning Knowledge and Reality*).
- ◆ The origins of neo-Fregean thinking are set out in the difficult Evans (1982) *The Varieties of Reference*.

The communicative or discourse failures of some victims of Alzheimer's are explored in Schwartz (1990). These sometimes include failures to grasp concepts holistically in webs of semantically related concepts, therein posing problems for the Davidsonian project of holistic interpretation.

References

- Bloor, D. (1997). *Wittgenstein on Rules and Institutions*. London: Routledge.
- Braude, S.E. (1997). A discursive account of multiple personality disorder. (Commentary on Gillett, 1997a) *Philosophy, Psychiatry, & Psychology*, 4(3): 223–226.

- Church, J. (2004). Social constructionist models: making order out of disorder—on the social construction of madness. In *The Philosophy of Psychiatry* (ed. J. Radden). Oxford: Oxford University Press, p. 11.
- Coulter, J. (1979). *The Social Construction of Mind*. London: Macmillan.
- Coulter, J. (1999). Discourse and mind. *Human Studies*, 22: 163–81.
- Davidson, D. (1980). *Essays on Actions and Events*. Oxford: Oxford University Press.
- Davidson, D. (1984). *Inquiries into Truth and Interpretation*. Oxford: Oxford University Press.
- Davidson, D. (1991). What is present to the mind? In *Consciousness* (ed. E. Villanueva). *Philosophical Issues*, Vol. 6: Ridgeview, pp. 197–213.
- Dennett, D. (1984). *Elbow Room*. Cambridge, MA: MIT Press.
- Dennett, D. (1987). True believers: the Intentional Strategy and why it works. In *The Intentional Stance*. Cambridge, MA: MIT Press, pp. 13–35.
- Dennett, D. (1991a). *Consciousness Explained*. Boston: Little Brown.
- Dennett, D. (1991b). Real patterns. *Journal of Philosophy*, 88: 27–51.
- Dennett, D.C. (1996). *Kinds of Minds*. London: Weidenfeld & Nicolson.
- Dummett, M. (1993). What is a theory of meaning I and II. In *The Seas of Language*. Oxford.
- Edwards, D. and Potter, J. (1992). *Discursive Psychology*. London: Sage.
- Evans, G. (1982). *The Varieties of Reference*. Oxford: Clarendon.
- Evrine, S. (1991). *Donald Davidson*. Oxford: Polity.
- Gillett, G. (1997a). A discursive account of multiple personality disorder. (Commentary by Braude, S.E. (1997, 223–226), and a response by Gillett, G. (1997, 227–230)) *Philosophy, Psychiatry, & Psychology*, 4(3): 213–222.
- Gillett, G. (1997b). Response to the Commentary on Gillett (1997a). *Philosophy, Psychiatry, & Psychology*, 4(3): 227–230.
- Harré, R. and Gillett, G. (1994). *The Discursive Mind*. London: Sage.
- Haugeland, J. (1993). Pattern and being. In *Dennett and his Critics* (ed. B. Dahlbom). Oxford: Blackwell, pp. 53–69.
- Kripke, S. (1982). *Wittgenstein on Rules and Private Language*. Oxford: Blackwell.
- LePore, E. and McLaughlin, B. (ed.) (1985). *Actions and Events*. Oxford: Blackwell.
- Luntley, M.O. (1999). *Contemporary Philosophy of Thought*. Oxford: Blackwell.
- McDowell, J. (1986). Singular thought and the extent of inner space. In *Subject, Thought and Context* (eds P. Pettit and J. McDowell). Oxford: Oxford University Press, pp. 137–158.
- McDowell, J. (1992). Meaning and intentionality in Wittgenstein's later philosophy. *Midwest Studies in Philosophy*, 17: 40–52.
- McDowell, J. (1999). In defence of modesty. In *Meaning Knowledge and Reality*. Cambridge, MA: Harvard.
- McGinn, M. (1999). *Wittgenstein's Philosophical Investigations*. Routledge.
- McMillan, J. (1999). Cognitive psychology and hermeneutics: two irreconcilable approaches? *Philosophy, Psychiatry, & Psychology*, 6: 255–258.
- Miller, A. (1998). *Philosophy of Language*. London: Routledge.
- Miller, A. and Wright, C. (ed.) (2002). *Rule-following and Meaning*. Chesham: Acumen.
- Mulhall, S. (1990). *On Being in the World*. London: Routledge.
- Quine, W.V.O. (1960). *Word and Object*. Cambridge, MA: MIT Press.
- Rosenthal, P.M. (1991). *The Nature of Mind*. Oxford: Oxford University Press.
- Rorty, R. (1993). Holism, intrinsicity and the ambition of transcendence. In *Dennett and his Critics* (ed. B. Dahlbom). Oxford: Blackwell, pp. 184–202.
- Sabat, S.R. (2001). *The Experience of Alzheimer's Disease*. Oxford: Blackwell.
- Sabat, S.R. and Harré, R. (1994). The Alzheimer's disease sufferer as a semiotic subject. *Philosophy Psychiatry and Psychology*, 1: 145–160.
- Schwartz, M. (ed) *Modular deficits in Alzheimer's-type dementia*. Cambridge MA: MIT Press.
- Thornton, T. (1998). *Wittgenstein on Language and Thought*. Edinburgh: EUP.
- Thornton, T. (2003). Psychopathology and two kinds of narrative accounts of the self. *Philosophy, Psychiatry, & Psychology*, 10: 361–367.
- Wittgenstein, L. (1953). *Philosophical Investigations*. Oxford: Basil Blackwell.
- Wright, C. (1987). On making up one's mind: Wittgenstein on intention. In *Logic, Philosophy of Science and Epistemology: Proceedings of the 11th international Wittgenstein symposium* (ed. P. Weingartner and G. Schurz). Vienna: Holder Pichler Tempsky, pp. 391–404.
- Wright, C. (1991). Wittgenstein's later philosophy of mind. In *Meaning Scepticism* (ed. Puhl), pp. 126–147.