

2

Divisions in folk psychology

This chapter sets out to challenge the unity assumptions and to establish the framework for a two-strand theory of mind. It highlights a number of divisions in folk psychology – distinctions we draw between types of belief and reasoning, and tensions in our thinking about the mind and mental explanation. These divisions do not compel us to abandon the unity assumptions; they can be explained away or dismissed as superficial. But, as we shall see, a two-strand theory can account for them in a particularly attractive way. The chapter is divided into three parts. The first looks at divisions in our view of belief, the second at related divisions in our view of reasoning, and the third at some deeper tensions in our view of the mind. Since the purpose of the chapter is to gather data, it will necessarily have a somewhat disjointed character, but connections will emerge as we go on, and by the end we shall have the outline of a tentative two-strand theory of mind.

I BELIEF

This part of the chapter reviews some distinctions we draw between different types of belief. There are some features common to all the types surveyed. They all have propositional content (or at any rate, *token* beliefs do; it is often argued that propositional content can be attributed to beliefs only as they are entertained in particular contexts); they all have mind-to-world direction of fit (they represent their contents as obtaining, rather than as to be made to obtain); and they all guide inference and action in a way that reflects their content and direction of fit – prompting actions and inferences that are rational in the light of them. These common features provide some grounds for the view that they are all variants of a single core state (what I called ‘the unity of belief assumption’). When we look more closely, however, the differences appear as marked

Divisions in folk psychology

as the similarities, and a serious challenge emerges to the unity of belief assumption.¹

1.1 Conscious versus non-conscious

The first distinction I want to mention is that between conscious and non-conscious beliefs. By *conscious* beliefs I mean ones that we are apt to entertain and act upon consciously. Non-conscious beliefs, on the other hand, are ones that influence our behaviour in an automatic, unreflective way, without being consciously entertained. For example, my behaviour when driving is guided by non-conscious beliefs about the rules of the road. It is worth stressing that to say that a belief is non-conscious, in this sense, is not to say that its possessor is not consciously aware of its existence. For example, noticing the way I place my feet as I walk down the street, I may consciously conclude that I have a non-conscious belief that it is dangerous to tread on the cracks in the pavement. That is to say, I can be conscious *of* a belief without the belief itself being conscious – that is, without its being apt to be consciously entertained and acted upon. (Consciously thinking that *I believe* that it is dangerous to tread on the cracks is different from consciously thinking that it *is* dangerous to tread on the cracks.)

Widespread acceptance of the existence of non-conscious mental states may owe something to the influence of Freudian psychoanalytic theory (though the ease with which elements of Freudian theory were absorbed into popular culture suggests that the seeds of the notion have been long present in folk psychology). However, the conception of the non-conscious mind invoked here is far less theoretically loaded than Freud's. The Freudian Unconscious is a collection of repressed memories and desires, often of a traumatic or sexual nature, which manifest themselves in pathological behaviours of various kinds. The non-conscious mind, as pictured here, is much more mundane. It consists for the most part of everyday beliefs and desires, such as my beliefs about the rules of the road, which are formed in the normal way and which help to shape normal behaviour. There is no implication that the states involved have been repressed, or that subject is unwilling to acknowledge their existence and

¹ For another approach to the ambiguity of belief, involving a fourfold classification, see Horst 1995.

Mind and Supermind

influence. (This is not to deny that some of our non-conscious beliefs and desires may conflict with our conscious ones, or that some may reflect hidden fears and anxieties. But these cases will be the exception rather than the rule.) The existence of non-conscious mental states of this anodyne kind is now widely acknowledged, and I shall assume that folk psychology embraces it.²

Now, in itself, the conscious/non-conscious distinction does not pose a significant challenge to the unity of belief assumption. Conscious and non-conscious beliefs might both belong to the same basic psychological kind, differing only in their possession, or lack, of some consciousness-conferring property – possibly a relational one (see, for example, Armstrong 1968, 1984; Carruthers 1996b; Rosenthal 1986, 1993). Indeed, token beliefs often seem to switch between conscious and non-conscious forms – sometimes being consciously entertained, sometimes influencing our behaviour non-consciously (again, my beliefs about the rules of the road are examples).

I shall be arguing that this is a mistake, and that conscious and non-conscious beliefs are of fundamentally different types. As we shall see, the conscious/non-conscious distinction aligns with a number of others, suggesting that conscious and non-conscious beliefs are differently constituted and belong to different systems. On this view, then, it is a mistake to think that a token belief can switch between conscious and non-conscious forms, sometimes operating at a conscious level and sometimes at a non-conscious one. In such cases, I shall argue, one has in fact two distinct token beliefs, of different types but similar content, which operate in different ways and at different levels. (Note, however, that I shall not be offering a theory of consciousness itself. I am interested in how conscious beliefs are constituted and how they function, not in what makes them conscious, and everything I say will be broadly compatible with all the various theories on the latter point.)

1.2 Occurrent versus standing-state

It is common to accept that beliefs can exist in two forms – both as dormant states of one's cognitive system and also as active mental events. At any moment, we all possess a huge number of beliefs which are not

² It is true that there is some philosophical resistance to the notion of non-conscious mentality (see, for example, Searle 1992), but I do not think that the folk share these worries.

Divisions in folk psychology

currently active in our minds. I believe – among many other things – that cider is made from apples, that my surname ends with an ‘h’, and that Tallahassee is the state capital of Florida. These are not things I think about much, but I believe them, and my belief would manifest itself in appropriate circumstances (if I were questioned, say). Beliefs like these are sometimes referred to as *standing-state* beliefs, and are contrasted with *occurrent* ones – that is, with episodes in which a belief is brought actively to mind. (The term ‘occurrent belief’ is not an everyday one, of course; indeed, there is no distinctive popular name for these episodes – ‘thought’ is perhaps the nearest, but it is not unambiguous.) We are often aware of experiencing such episodes. For example, just now I was reflecting that the weather is unseasonably warm and thinking that it would be wise to turn the thermostat down. Episodes like this can occur spontaneously without any effort on our part – thoughts just *strike us* or *pop into our heads* – but we also can set out to induce them deliberately, as when we *rack our brains* for the answer to a question. Occurrent thoughts – at least when conscious – invariably occur serially, and they often form coherent sequences, linked by bonds of association or justification. Much of our conscious mental life consists of such trains of thought. This is the so-called ‘stream of consciousness’ which some novelists have tried to reproduce.³

Most writers on belief have recognized the existence of both occurrent and standing-state beliefs, though not all have given them equal weight in their theorizing. Early modern philosophers tended to concentrate on occurrent belief and to neglect the standing-state variety. Twentieth-century behaviourists, by contrast, switched the focus to standing-state belief and largely ignored the occurrent form. In part, this reflected a difference of theoretical interest – in one case in the phenomenology of belief, in the other in its role in the explanation of action – but an adequate account of belief should accommodate both aspects.⁴ The dominant contemporary view – the representational theory of mind – sees the two varieties as different aspects of the same state, differing in their level of activity. The theory identifies standing-state beliefs with stored representations, and occurrent beliefs with activations of these representations,

³ For discussion of the nature of occurrent thought and an argument for its distinctness from other kinds of mental state, see Swinburne 1985.

⁴ Henry Price characterizes the traditional view as the Occurrence Analysis and the behaviourist one as the Dispositional Analysis, and argues that elements of both are needed (Price 1969). For an account of belief which combines both phenomenological and action-based criteria, see Braithwaite 1932–3.

Mind and Supermind

preparatory to their employment in reasoning and decision-making (see, for example, Fodor 1987). On this view, then, occurrent belief again occupies centre stage, with standing-state beliefs remaining idle until activated in occurrent form.

This view has an intuitive appeal. It does often seem as if occurrent activation is required for a belief to play a role in reasoning and decision-making. I am driving to work as usual. Suddenly, it occurs to me that roadworks are due to start today, and I decide to take a different route. I do so, moreover – or so it seems to me – precisely *because* the thought about the roadworks occurred to me. Had it not done so, I would not have changed course (other things being equal, that is). The belief's emerging in occurrent thought was necessary for it to play a role in my decision-making. Or think about absent-mindedness. The light bulb blows and I go off in search of a new one. When I return to the room, my hand again goes to the light switch. Why? It seems that I had forgotten that the bulb was blown. Yet surely my memory is not that bad. I had not *ceased to believe* that the bulb was blown and would immediately have avowed that belief if questioned. The belief, it seems, was stored in my memory, but somehow failed to influence my behaviour. The natural way of explaining this would be to say that it *failed to occur* to me – that is to say, that it failed to influence how I behaved because it failed to become occurrent at an appropriate moment.⁵

We hold, then, that some standing-state beliefs require activation in occurrent form in order to influence action. But it is not clear that we think that all of them do. We certainly do not suppose that all beliefs require activation as *conscious* occurrent thoughts. We accept that beliefs can influence action non-consciously. And we allow that animals act on their beliefs – without, I think, thereby committing ourselves to the view that they have conscious occurrent thoughts. (As Malcolm notes, it would sound funny to say of an animal that a thought *occurred to him*, or *struck him*, or *went through his mind*; see Malcolm 1973.⁶) Of course, it is possible that in these cases the relevant beliefs are activated as *non-conscious* occurrent thoughts, but, as we shall see, it is doubtful that there is any folk

⁵ For further defence of this view, see Goldman 1970, ch. 4.

⁶ This indicates, Malcolm goes on, that *having thoughts* is not the paradigmatic form of mental activity. For we find it very natural to speak of animals *thinking* and to explain their behaviour by reference to what they think. Malcolm concludes that there is no single paradigm or prototype of thinking and suggests that it was by treating *having thoughts* as the paradigm form that Descartes was led to deny that animals can think.

Divisions in folk psychology

commitment to this view. That is to say, occurrent belief, as we commonly conceive of it, may be unique to the conscious mind, and a behaviourist perspective may be more appropriate for the non-conscious mind. I shall say more about this in the next part of this chapter.⁷

With the distinctions between standing-state and occurrent belief in place, I can now restate more clearly the proposal I am making. I am suggesting that there are two types of standing-state belief: those of the first type receive activation as conscious occurrent thoughts and influence action at a conscious level, and those of the second do not receive conscious occurrent activation and influence action non-consciously. Henceforth, when I talk of *conscious* and *non-conscious* beliefs it is these two types of standing-state belief I shall mean. (Of course, there is a sense in which *all* standing-state beliefs are non-conscious, in virtue of the fact that they are not currently present to consciousness, but I am using the term to denote availability to conscious thought rather than actual presence.) Note that a standing-state belief counts as conscious only if it is apt to be activated as a conscious occurrent thought *with the same content*, and that we can entertain conscious occurrent thoughts *about* our non-conscious standing-state beliefs without thereby rendering those beliefs conscious. So, for example, I can entertain the conscious occurrent thought that I have a non-conscious standing-state belief that it is dangerous to tread on the cracks in the pavement without thereby coming to have a conscious standing-state belief that it is dangerous to tread on the cracks in the pavement.

1.3 Flat-out versus partial

We often speak of belief as a binary, or flat-out, state, which is either categorically present or categorically absent. So, for example, I believe that my car is green but do not believe that it has an automatic gearbox. A binary view of belief is also implicit in much of our reasoning, which frequently operates upon unqualified propositional attitudes. We reason, for example, that we want a beer, that if we go to the fridge we can get a beer, and thus that a trip to the fridge is called for. This is not the whole picture, however. For we also speak of having degrees of confidence, or 'partial beliefs', which are continuously variable. So, for example, I am very confident (though not certain) that tap water is safe to drink, somewhat

⁷ It is worth stressing that not all occurrent thoughts are activations of previously formed beliefs; some involve the formation of new ones, while others are idle speculations or fantasies. I shall say more about these other forms of occurrent thought later.

Mind and Supermind

less confident that my car is in good working order, and still less confident that it will not rain today. Degrees of confidence are sometimes referred to as *subjective probability assignments*, and it can be argued that rational decision-making should be sensitive to them, in the way prescribed by Bayesian decision theory. (I shall say more about this shortly.) The easiest way of ascertaining what degree of confidence a person has in various propositions is to offer them bets on their truth; if they are rational and the payoffs are linear, then their betting behaviour will vary in accordance with their degrees of confidence.

So here we have another division – between a qualitative, flat-out form of belief and a quantitative, partial form. This is not, of course, to say that the two forms are fundamentally distinct. We might maintain that one of them is the core form of belief and the other a derivative or subspecies of it. There are two options here, depending on whether we take partial or flat-out belief to be the core state. Neither is particularly attractive, however. The former option – taking partial belief as core – has been widely canvassed, and I shall consider it in detail in the next chapter. Here I want to deal with the latter – the view that flat-out belief is the core state and partial belief the derivative.

What are the possibilities here? One option, advocated by Gilbert Harman, is to suppose that when we talk of degree of belief we are referring to how strongly held our *flat-out* beliefs are, where this is a matter of how hard it would be for us to give them up (Harman 1986, ch. 3). So, one either believes a proposition or does not believe it; but if one does, then one does so with a certain degree of strength or attachment. (Harman stresses that this need not involve making an explicit assessment of how important the belief is to us; our degree of attachment to it may be implicit in the way we reason – the more attached to it we are, the more powerful the reasons needed to get us to abandon it.) Now, I think that it is quite right to say that flat-out beliefs can be held with different degrees of attachment, but it is implausible to identify degrees of confidence with these degrees of attachment. For we can have a degree of confidence in a proposition without having a flat-out belief in it at all. I am fairly confident that it will not rain tomorrow, but I do not believe flat-out that it will not. Indeed, the claim that degrees of confidence require flat-out belief leads to absurdity. For according to Bayesian principles, rational agents will entertain some degree of confidence in every proposition of whose falsity they are not completely certain – including pairs that are contradictory. Yet it is absurd to say that a rational agent will have a

Divisions in folk psychology

flat-out belief in every contingent proposition. This option is unattractive, then.

A second option is to hold that when we talk of degrees of confidence what we are really referring to is flat-out beliefs in objective probabilities – where to say that an event has a certain objective probability is to say something about the frequency with which events of that type happen. So, for example, when we say that someone is 50 per cent confident that a coin toss will come up heads, what we mean is that they believe flat-out that the objective probability of its coming up heads is 0.5, where this in turn means that they believe flat-out that, if tossed repeatedly, the coin would come up heads 50 per cent of the time. This view is unattractive, however, for two reasons. First, it makes the ability to entertain beliefs about objective probabilities and frequencies a prerequisite for having partial beliefs. And this is implausible. We want to say that people unfamiliar with those notions can nonetheless have degrees of confidence. (This is not to say that we never form flat-out beliefs about objective probabilities, just that our degrees of confidence cannot be identified with them.) Secondly, on the proposed view it would follow that we cannot have degrees of confidence in single events, since single events do not have objective probabilities, understood as frequencies. It is meaningless to talk of the frequency of a single event. Yet we do have degrees of confidence in single events – for example, I am fairly confident that my friend will call this evening to return the book I lent her. So degrees of confidence cannot be construed as flat-out beliefs in objective probabilities.

A final option is to identify degrees of confidence with flat-out beliefs about one's own behavioural dispositions – say, about how willing one would be to bet on various outcomes. So, for example, we might say that to be 50 per cent confident that one's car is in good working order is to have the flat-out belief that one would be willing to bet on the car's being in good working order at odds of evens or better. This proposal is also unattractive, however. For, as with the previous one, it overestimates the intellectual requirements for having partial beliefs. One can have a degree of confidence in something without having beliefs about one's betting behaviour, or, indeed, without understanding how betting works. A similar objection will hold, I think, for any other analysis of degrees of confidence in terms of flat-out beliefs about one's own behavioural dispositions.

On a first pass, then, the division between flat-out and partial belief stands up well. Of course, this still leaves us with a question about the nature

Mind and Supermind

of partial belief. What exactly is it to have a certain degree of confidence in something? I shall return to this question later in the chapter.

1.4 Active versus passive

There is a long tradition in philosophy of maintaining that beliefs can be actively formed – that we have the power to decide what attitude to take towards a proposition, through an act of deliberate *judgement* or *assent*. The idea is that we can consider a proposition, reflect upon the evidence for and against it, and then decide whether or not to accept it as an object of belief. In the past, many philosophers took it for granted that we have this power, and many would have identified what I have been calling *occurrent beliefs* with episodes in which a thought is entertained prior to assent or rejection.⁸ Yet many contemporary philosophers deny that we have a power of active judgement, and insist that belief formation is always passive. They concede, of course, that we can indirectly influence what we believe – for example, by practising autosuggestion or by selectively focusing on favourable evidence – but deny that we can form beliefs directly, by one-off acts of judgement.⁹

Both points of view have some plausibility. It is undoubtedly true that much belief formation is passive. One has only to think of those beliefs that derive from perception and memory. In such cases, belief is forced upon us without any effort or choice on our part. I assume that non-conscious beliefs, too, are passively formed. Sometimes, however, we seem to take a more active role in belief formation. We frequently talk of *deciding*, *judging*, or *making up our minds* that something is true – and we speak of these episodes as datable, one-off actions which directly produce belief.

Talk of *making up our minds* is particularly suggestive here. Of course, sometimes, when we speak of a person having made up their mind, what we mean is that they have made a decision to do something – that they have formed an intention, not a belief. But we also speak of making up

⁸ The most famous defence of the freedom of assent is in book 4 of Descartes's *Meditations* (Descartes 1984, vol. I). See also his *Principles of Philosophy* 1, 39 (Descartes 1984, vol. II).

⁹ These writers would agree with Hume's assertion that belief 'depends not on the will, but must arise from certain determinate causes and principles, of which we are not masters' (Hume, 1739/1888, p. 624), though few would endorse his reason for it – that belief is a feeling or sentiment (but see Cohen 1992).

Divisions in folk psychology

our minds about matters of fact – about the truth of a theory, say, or the safety of a course of action. For example, suppose that there is controversy about the safety of eating beef: a deadly disease is rife among cattle, and some scientists claim that it can be transmitted to humans through beef products, though others insist that adequate safety measures are in place. Is it safe to eat beef?¹⁰ The evidence is inconclusive, but we need to take a view. So, having reflected on the evidence and the risks, we make up our minds on the matter. We also speak of making up our minds about what we *want*, and often urge people – children especially – to do it. These phenomena have received little attention from contemporary philosophers of mind. Yet, as Annette Baier emphasizes in one of the few discussions of the subject, they have some distinctive features (Baier 1979). (Baier focuses on the process of *changing* one's mind, but her observations apply equally to that of making it up, which she treats as the original activity of which a change of mind is the revision.) A change of mind, Baier argues, is a special kind of cognitive process, distinct from the routine acquisition or updating of information. If I think that the ice will support me, and discover from bitter experience that it will not, then I can be said to have *learned better*, but not to have *changed my mind*. Changes of mind, Baier goes on, are not forced upon us, by either external or internal influences, but are the product of free reflective judgement. They follow upon a re-evaluation of our options, perhaps in the light of new or previously neglected information, and they involve a considered judgement about the appropriateness of some action or attitude. The moral of Baier's analysis is that change of mind is a genuinely personal activity – it is something one *does*, not something that happens to one, and it requires attitudes and skills of some sophistication.

Of course, to say that making up, or changing, one's mind involves active reflection is not to say that it involves active judgement. Perhaps all we actively initiate is the reflection; we just have to wait and see whether it subsequently produces a new belief. That is a possible view of the matter – but not, I think, an attractive one. The point of making up, or changing, one's mind about a topic is precisely to *settle*, or *resettle*, what one thinks about it. And one does not achieve this simply by pursuing reflection in

¹⁰ Safety is, of course, a relative matter. In asking whether it is safe to eat beef, I mean safe by whatever standard one adopts for matters of this kind. With the presence of BSE (or 'mad cow disease') in the British beef herd, and the apparent emergence of a human form of the disease, this question was one that faced many British consumers in the mid 1990s.

Mind and Supermind

the hope that it will generate a robust conviction. That is a good way never to arrive at a settled view of anything. To make up one's mind, one needs to be able, not only to initiate and sustain reflection, but also to *foreclose* on it – to cease pondering and adopt a definite view of the matter. We must, as Henry Price puts it, 'come off the fence on one side . . . prefer or plump for one of the alternatives, accept it or commit ourselves to it, and reject the others' (Price 1969, p. 206). This process, Price suggests, is like the resolution of a conflict ('making up' a previously 'divided' mind) and is analogous to the decision-making that terminates practical reflection (p. 206).¹¹ On this view, then, changing or making up one's mind involves two elements: a reflective re-evaluation of one's options, and a kind of free, creative judgement. It involves looking *and then leaping*.

This suggests, then, that we need to recognize the existence of two broad kinds of belief formation and revision: one passive and unreflective, the other involving personal reflection and active assent.

1.5 Language-involving versus not language-involving

It is sometimes claimed that we can *think in* natural language – that natural language can act, not only as a medium for the expression of thoughts, but also as a vehicle of thought itself (Bickerton 1990, 1995; Carruthers 1996b; Dennett 1991a; Harman 1973). This view has a strong intuitive appeal. We often seem to form or recall a belief in the act of verbally articulating it to ourselves. Think, for example, of looking out of the window, seeing the louring clouds, and saying to oneself, 'It's going to rain.' Here, it seems, one does not form the belief and then articulate it (why would one do that?). Rather, one forms the belief in the act of articulating it; the linguistic action is partially constitutive of the thought. Many people claim that much of their conscious thinking takes place in this way, in the form of a self-directed sub-vocalized monologue. There are also theoretical reasons for believing that some kinds of thought involve natural language, and a powerful argument can be run for the view that conscious propositional thinking is language-based (Carruthers 1996b, 1998). But, of course, it is implausible to suppose that *all* thought involves natural language. It is hard to deny that animals and pre-linguistic infants can think, and there is no pre-theoretical reason to suppose that our non-conscious thoughts involve

¹¹ Price makes these remarks in the course of outlining a view ('the Occurrence Analysis', as he calls it) which he does not himself fully share; but he gives every impression of endorsing them.

Divisions in folk psychology

natural language. So here, again, there is a division between two types of thought.

(It is worth stressing that the distinction just made concerns *medium of representation*. To say that a thought is not language-involving is to say that it does not employ natural language as a medium of representation. Such a thought may, however, be dependent on language in another way. There are beliefs which we would never form without language, simply because the concepts involved are ones that can only be *acquired* through language. Beliefs about black holes, for example, are language-dependent in this way, since without language we would never acquire the concept of a black hole. I shall not be concerned here with this sort of language dependency, and shall assume that it is independent of the previous sort—that a thought may be language-dependent in this sense without being language-involving and vice versa.)

1.6 Some links

We have looked at some divisions in our view of belief. They are a varied bunch, relating to consciousness, activation level, degree, mode of formation, and medium of representation. Already some links are emerging, though. Conscious beliefs are activated in episodes of occurrent thought, and such episodes often seem to involve acts of judgement or assent. Assent, in turn, introduces a qualitative attitude: one either assents to a proposition or does not; there is no halfway house (Price 1969, p. 207). (Of course, one might assent to an estimate of probability, but then it is the content of the attitude that is qualified, not the attitude itself.) And our conscious thoughts and judgements are often framed in natural language (though they may not always be so; I shall say more later about the extent of language involvement in conscious thought). Non-conscious beliefs, on the other hand, are different. They are not actively formed, and there is no pre-theoretical reason to think that they require occurrent activation or involve natural language. It is plausible, too, to locate our partial beliefs at the non-conscious level; our degrees of confidence are not matters of immediate conscious awareness and reveal themselves most clearly in our behaviour. So we can tentatively distinguish two types of belief: one which is conscious, subject to occurrent activation, flat-out, capable of being actively formed, often language-involving, and, consequently, unique to humans and other language users; and another which is non-conscious, possibly not subject to occurrent activation, partial, passively

Mind and Supermind

formed, probably non-verbal, and common to both humans and animals. We have, then, the beginnings of a two-strand view of belief.

The reader may ask why we should regard these two kinds of belief as distinct psychological states, rather than as varieties of a single generic type. After all, I call them both *beliefs*, so I must be assuming the coherence of some broader notion of belief which encompasses both. My response is to distinguish levels of description. At a very broad level of description, it may be appropriate to stress the similarities between the two kinds of belief and to group them together for explanatory and predictive purposes. That is what folk psychology does, and for its purposes such a classification is generally quite adequate. But at a finer level of description, it is important to distinguish them and to recognize that neither can be regarded as a subspecies of the other. Moreover, it is at this level of description that many philosophical debates about belief are pitched. Do beliefs require occurrent activation? Are they graded or ungraded? Can they be actively formed? Do they involve language? These are questions we need to resolve if we are to integrate the folk concept into a serious science of the mind, but they cannot be resolved so long as we stick with the inclusive folk classification, since they have no uniform answer. To repeat a point made earlier: the pressure to distinguish the two kinds of belief comes from the demands of integrationism.

Finally, a word about desire. Although I have focused on belief, very similar considerations apply to desire. It, too, appears to have two forms – one conscious, apt to be occurrently activated, active, flat-out, and frequently language-involving; the other with the opposite properties. I shall assume, then, that each form of belief is associated with a complementary form of desire, giving us a two-strand view of both states.

2 REASONING

I shall now move on to challenge the unity of processing assumption – the assumption that we have a single, unified, reasoning system. Again I shall review a number of common-sense distinctions and suggest that, collectively, they motivate a two-strand approach. The claim that we should reject the unity of processing assumption is not without precedent. A number of psychologists have proposed *dual-process* theories of reasoning, distinguishing a slow, serial, rule-based, conscious system and a fast, parallel, associative, non-conscious one (see Evans and Over 1996; Wason

Divisions in folk psychology

and Evans 1975; Sloman 1996; Stanovich 1999).¹² These theories are motivated by experimental data on reasoning and rationality, but, as we shall see, a similar picture is latent within folk-psychological discourse itself.

2.1 Conscious versus non-conscious

I have already introduced the idea that some of our reasoning is non-conscious. Everyday experience, I think, strongly supports this view. We can perform many complex activities without conscious thought. Think of driving, for example. Controlling a car, following a route, anticipating the behaviour of other road-users – these are difficult tasks requiring considerable intelligence. Yet we often perform them without any conscious thought at all. Or consider chess. Skilled chess players can evaluate hugely complex positions very quickly, with little conscious thought. Typically, they will consciously assess only a few of the most promising strategies, ignoring the many thousands of others available to them. Again, it seems, much rapid and complex non-conscious processing must be involved in weeding out the unpromising options and selecting a few of the best for conscious evaluation.

The existence of non-conscious reasoning presents an immediate challenge to the unity of processing assumption. For conscious and non-conscious reasoning can proceed independently, on different topics – strongly suggesting that the two are supported by distinct systems. (Think of driving again: one can drive to work, non-consciously making all the required calculations, while one's conscious mind is wholly occupied with other matters.) This conclusion is reinforced by reflection on the role of conscious thought in the guidance of action. Suppose that there were just a single reasoning system, whose processes were sometimes conscious and sometimes not. Then a belief would come to influence action by being taken as an input to this system, and it would not much matter whether this happened consciously or non-consciously. That is to say, consciousness would be an optional extra, which made no direct difference to a belief's causal role (though it might, of course, make an *indirect* difference,

¹² Paul Smolensky sketches a similar picture from the perspective of cognitive science, distinguishing a non-conscious intuitive processor, which operates on connectionist principles, and a conscious rule interpreter, which executes rules formulated in natural language (Smolensky 1988).

Mind and Supermind

by making the subject aware of its existence). And this is counter-intuitive. Introspection suggests that the conscious status of an occurrent thought is as important to its causal role as its activity level. Recall my thought about the roadworks, which led me to deviate from my normal route to work. It was – or so it seemed to me – precisely because I *consciously* recalled that the roadworks were due to start that I changed course. If I had not done so, I would not have acted upon that information – at any rate, not in that way, at that time. Consciousness was not an optional extra, but crucial to the belief's efficacy. This suggests that conscious activation involves a distinct pathway to behavioural influence, independent of the non-conscious one.

It seems likely, then, that the conscious and non-conscious reasoning systems are distinct. It is compatible with this, however, that both systems are similarly constituted. Processing could be of a unitary *kind*, even if it proceeds in more than one *strand*, and in this weaker sense the unity of processing assumption could be sustained. In what follows I am going to question whether even this weaker sort of unity holds. (Henceforth, references to the unity of processing assumption are to this weaker claim.) As with belief, the division between conscious and non-conscious reasoning aligns with various others, suggesting that the two systems are differently constituted and that the states directly available to the one are not directly available to the other.

2.2 Explicit versus non-explicit

Conscious reasoning typically involves entertaining sequences of explicit propositional thoughts – occurrent beliefs and desires – which form chains of sound or probable inference. I shall refer to reasoning of this kind as *explicit* reasoning and shall contrast it with *non-explicit* reasoning. Note that 'explicit' here does not imply 'conscious': it is possible that non-conscious reasoning involves explicit propositional thoughts, too (though I shall be questioning whether it does in fact do so). 'Non-explicit' is a blanket term for any mental processing which does not involve explicit propositional thoughts.

Now, conscious reasoning is explicit, but what of the non-conscious variety? Those who subscribe to the unity of processing assumption will say that it too is explicit – that non-conscious thought processes involve sequences of non-conscious occurrent beliefs and desires. It is not obvious, however, that there is any folk commitment to this view. Indeed,

Divisions in folk psychology

it seems rather counter-intuitive. For one thing, it is doubtful that there would be time for all the necessary non-conscious occurrent thoughts to occur. Think, for example, of all the calculations needed in order to guide a car safely from one's home to one's workplace. One needs to calculate how to steer in order to avoid obstacles and follow the desired route; how fast to travel, given the traffic laws, road conditions, and time available; when to overtake; when to change gear; when and how sharply to brake; and so on. This is a massive computational burden, and it is hard to see how it could be discharged if every relevant belief and desire had to be activated in occurrent form. At any rate, it does not seem to be part of the folk view that such activation *must* occur. (As Andy Clark points out, skilled drivers do not *expect* their actions to be the product of sequences of discrete inferential steps; Clark 1993b, p. 230.) Or think about the expert chess player. It would require huge sequences of explicit propositional operations to arrive at judgements of the sort routinely made by skilled chess players – sequences which only the most powerful computers can execute. It seems more likely that the processes involved are non-explicit. Or, finally, consider inference. A friend points to a parked car which has its headlights on and says, 'The owner of that car will be annoyed with himself.' I immediately understand the remark; yet its comprehension depends on a host of beliefs about the nature and function of cars and the beliefs and habits of car drivers. And, again, it is doubtful that there would be time for all of these to be individually accessed and activated.

It may be objected that I am adopting too strict an interpretation of the claim that non-conscious reasoning is explicit. Not *all* the beliefs and desires involved in an episode of explicit reasoning need be activated as occurrent thoughts. After all, even conscious reasoning often depends on suppressed premises and background assumptions. We think, 'Looks like rain – better take the umbrella.' We do not add 'Rain is unpleasant', 'Umbrellas protect against rain' – though, of course, we believe these things and would not arrive at the conclusion if we did not. And, it might be suggested, the same may go for non-conscious reasoning – many of the premises involved may be implicit. This is a fair point, but it does not do much to support the claim that non-conscious reasoning is explicit. Non-conscious reasoning often involves making extremely complex assessments, drawing on many different factors (again, think of driving or chess-playing), and it seems unlikely that such assessments could be plausibly reconstructed as sequences of occurrent thoughts, even if some

Mind and Supermind

of the premises were suppressed. Besides, the fact that conscious reasoning depends on suppressed premises and background assumptions hardly tends to support the view that all reasoning is explicit – rather the opposite. If beliefs can influence our reasoning without being activated as occurrent thoughts, then why are *any* of them activated in that way? Why is not all reasoning non-explicit?¹³

So we have a tension. The folk are committed to the view that conscious reasoning is explicit, but are not committed to the view that non-conscious reasoning is. So they are not committed to the unity of processing assumption, and should be at least receptive to the idea that there are two strands or levels of reasoning – one conscious and explicit, the other non-conscious and perhaps involving non-explicit processes. This in turn tends to confirm that we have two kinds of belief: conscious beliefs, which require activation as occurrent thoughts, and non-conscious ones, which may not.

Let me pause to add a qualification to the picture that is emerging. I have suggested that conscious reasoning forms a distinct strand of cognition, and that the beliefs which figure in it require activation in occurrent form. This claim needs qualifying, however. For, as I noted above, conscious reasoning often depends on suppressed premises and background assumptions which are not explicitly activated. (By background assumptions I mean beliefs which influence the outcome of our reasoning without figuring as premises in it. For example, the assumption that it is safe to go out at night may influence my practical reasoning about what to do tonight, affecting which options are considered and which conclusions embraced.¹⁴) I shall say that beliefs which form suppressed premises and background assumptions in a reasoning episode are *implicitly active* in it, and I want to make two points about them. First, we should not think of these beliefs as belonging to a different strand of mind from their explicit counterparts. The distinction between the two does *not* correspond to that between the

¹³ Arthur Walker has defended the view that reasoning involves occurrent thoughts against the rival view that it consists in transitions between standing-state beliefs (Walker 1985). Only on the former view, Walker argues, can we make sense of the idea that a person may *re-infer* an already accepted conclusion from new evidence. While I find this persuasive as an argument for the view that *some* of our reasoning must be explicit, I do not think that it supports the view that *all* of it must be. (Walker does not consider non-conscious reasoning and seems to view inference as a conscious phenomenon with a distinctive phenomenal aspect; p. 207.)

¹⁴ Michael Bratman emphasizes the necessity of making background assumptions like these in one's practical reasoning (Bratman 1987, 1992).

Divisions in folk psychology

two strands of reasoning I have tentatively identified. Rather it is *intrinsic* to the conscious explicit strand. A belief counts as implicitly active only relative to an episode of explicit reasoning – it is one that is involved in the episode without being explicitly involved in it. If non-conscious reasoning is non-explicit, then we cannot make an explicit/implicit distinction for the beliefs involved in it – in a sense they are all implicit. (Another reason for regarding suppressed premises and background assumptions as belonging to the conscious level is that, like conscious beliefs in general, they are *flat-out* ones. They are things we *take for granted* in our reasoning, and taking for granted is an all-or-nothing attitude. Even probabilistic reasoning, when explicit, requires flat-out background assumptions; see Lance 1995.) Secondly, although it is not true to say that *all* the beliefs involved in conscious reasoning require occurrent activation, it remains true, I think, that all the *non-obvious* ones do. Suppressed premises and background assumptions are things we take for granted in our reasoning – they define the normal background to it. And any belief which does not form part of the normal background – which is, let us say, *epistemically salient* – will typically require occurrent activation.

2.3 Classical versus probabilistic

It is plausible to think that rational decision-making should be sensitive to considerations of probability. Suppose that I am trying to decide what to do on my afternoon off: play squash, go for a walk in the country, do the shopping, take the car in for servicing, and so on. Which is most desirable? In each case the desirability of the action will vary depending on what background conditions obtain. If my knee injury is not completely healed, then it would be better to avoid squash; if it rains this afternoon, then a walk in the country would not be fun; and so on. So my decision should reflect how probable I think it is that each of these conditions obtains and how desirable I find the various outcomes contingent upon them. (These values are known as my subjective probabilities and desirabilities – *subjective* since they reflect my idea of what is probable and desirable, rather than some independent measure of those things.) These intuitions can be developed into a full-blown probabilistic decision theory – Bayesian decision theory. In brief, the procedure involves taking each candidate action in turn and calculating its *weighted desirability* relative to each possible background condition. This is given by multiplying the desirability of the outcome the action would have if the background condition obtained

Mind and Supermind

by the probability that the condition does obtain. Summing these values for each background condition gives the overall estimated desirability of the action. The optimum action is the one with the highest estimated desirability.¹⁵ There is also a well-developed probabilistic logic of inductive inference – Bayesian confirmation theory – which tells us how to adjust our confidence assignments in the light of new evidence.

Now, these theories are intended primarily as normative, not descriptive, ones. Nevertheless, much of our reasoning does yield to description in Bayesian terms. In particular, our decision-making can often be interpreted as the upshot of probabilistic reasoning, sensitive to the sort of factors mentioned. Indeed, it can be shown that any agent whose preferences satisfy certain intuitively reasonable conditions can be interpreted as assigning degrees of probability and desirability to relevant sets of propositions and outcomes and as maximizing estimated desirability relative to those assignments. (Demonstrations of this are called *representation theorems*; see Ramsey 1926; Savage 1972; and, for surveys, Eells 1982; Fishburn 1981.) Animal behaviour, too, can often be interpreted in this way (Jeffrey 1985).

It may be objected that to say that we can be *represented* as engaging in Bayesian reasoning is not to say that we *do* engage in it: representation theorems do not show that Bayesian decision theory characterizes internal psychological reality (Goldman 1986, p. 327). Indeed, there is reason to think that the human brain does not perform calculations of probability and desirability, but relies instead on ‘fast and frugal heuristics’ which generate responses that are quite good enough for most everyday purposes (see, for example, Gigerenzer et al. 1999). I am going to set this objection aside for a moment. How serious it is depends on what the function of a theory of mind is, and I shall return to it later in this chapter, when we have looked at that question.

Grant for the moment that at least some of our decision-making can legitimately be characterized in Bayesian terms. An odd consequence follows. For our *conscious* reasoning very rarely takes a Bayesian form. We generally prefer to reason from unqualified premises to unqualified conclusions, employing classical inference schemata, such as the practical syllogism. And we find it very hard to identify our own assignments of

¹⁵ For detailed explication and defence of Bayesian decision theory, see Jeffrey 1983; Kaplan 1996; Savage 1972. It is common to refer to desirabilities as ‘utilities’ and estimated desirability as ‘expected utility’. The terminology used here (which is borrowed from Jeffrey 1983) seems to me more natural.

Divisions in folk psychology

probability and desirability, as revealed in our choices. So we have a puzzle. There are two incompatible ways of explaining a subject's decisions and inferences: by reference to their conscious classical reasoning and their professed beliefs and desires, and by reference to non-conscious Bayesian calculations involving assignments of probability and desirability of which the subject is unaware.¹⁶

How should we respond to this puzzle? One option – proposed by Ronald de Sousa – is to think of the two kinds of explanation as characterizing different *levels* of mental processing. There is a level of non-conscious non-verbal deliberation, de Sousa suggests, which is common to humans and animals and whose workings can be characterized in Bayesian terms, and there is a level of conscious verbalized reasoning which is found only in humans and which operates according to classical principles (de Sousa 1971, pp. 57–8). Here, then, is another motive for questioning the unity of processing assumption.

2.4 Active versus passive

I suggested earlier that beliefs can be actively formed, and I now want to suggest that they can be actively processed too. Some reasoning processes, I suggest, are intentional actions, initiated and controlled at a personal level. This is most obvious in cases where we employ some explicit inferential procedure – constructing a syllogism, say, or writing out a long division. In such cases the overt actions involved can be thought of as constituting a larger inferential action – *making a calculation* or *deriving a conclusion* – which is under fully personal control. We can do similar things in our heads, articulating an argument in inner speech or visualizing the steps of a mathematical calculation. This is not all, however. Even when no explicit procedures are employed, it can still be appropriate to think of an inference as intentional. I am pondering the car's innards, trying to identify a particular component. 'That is the cylinder head and that is the air inlet', I mutter to myself, 'but what is *this*?' – staring at the component in question and furrowing my brow. Suddenly the answer comes to me:

¹⁶ As Hempel notes, this gives a 'peculiar twist' to the idea of rational action:

though . . . subjects make their choices in clearly structured decision situations, with full opportunity for antecedent deliberation and even calculation, they act rationally (in a precisely defined quantitative sense) relative to subjective probabilities and utilities which they do not know, and which, therefore, they cannot take into account in their deliberations. (Hempel 1965, p. 483, quoted in de Sousa 1971, p. 57)

Mind and Supermind

'It must be the fuel lead.' Here I do not employ any explicit procedure to arrive at the answer and I have no conscious awareness of how I get there. But still, I suggest, getting there involves personal activity on my part, though it is hard to say exactly what this activity is. (We would simply say that I was *thinking*, or *trying to work out*, what the component was. I shall offer a more illuminating characterization in chapter 4.) In claiming that some reasoning episodes are intentional, I am, of course, supposing that they have belief/desire explanations, and the reader may ask what the motivating beliefs and desires for a reasoning episode might be. The answer, I think, is simple: the desire to find a solution to some problem and the belief that doing *this* – writing out a syllogism, running through an argument in inner speech, or just 'thinking' – is a way to get one. These beliefs and desires will, I assume, usually be non-conscious.

Of course, it would be wrong to suppose that all, or even most, of our reasoning involves personal activity, even of the inchoate kind just described. Sometimes a solution pops into our heads unbidden long after we have consciously given up trying to find it. And, of course, we have no direct control over our non-conscious reasoning. (It is true that we tend to use active verbs for non-conscious inferences. We say such things as 'I braked because I realized that the other car was going to pull out', even though the action was not preceded by any conscious inferential activity. But these locutions should not be taken literally – compare the way we speak of *digesting* our food. In such cases, I suggest, we are inferring the non-conscious reasoning that led to the action, and assimilating it to the pattern of our conscious active deliberation.) So here is another division – between reasoning that is intentional and reasoning that is not.

2.5 Language-driven versus not language-driven

The final distinction I want to mention concerns the role of natural language in reasoning. I claimed earlier that natural language can serve as a medium for the *representation* of thoughts, but it is plausible to think that it can serve as a medium of *inference*, too. Some reasoning processes, it seems, constitutively involve the manipulation of natural-language sentences – written, vocalized, or, most often, articulated in inner speech. As I mentioned in the previous section, we can use language to perform explicit inferential operations, as in the construction of syllogisms. There are also other, less formal, examples of language-based reasoning. We often reason

Divisions in folk psychology

things out in the course of conversation with a friend or in interior monologue with ourselves. In such cases, it seems, we are not simply recapitulating reasoning that has already been conducted in some inner medium, but conducting the reasoning in the course of articulating it: the linguistic activities *implement* the reasoning process, carrying it forward and shaping its direction. (Since language use is an intentional activity, this observation lends support to the earlier suggestion that some reasoning is active.)

It seems, then, that some of our reasoning is language-driven. It is implausible, however, to suppose that *all* of it is. Animals, I take it, are capable of reasoning, as are people with severe aphasia (Varley 1998). Nor is there any obvious reason to think that non-conscious reasoning is language-driven. It certainly does not involve inner speech, and while it might involve sub-personal linguistic processes, there is no pre-theoretical reason to think that it does. Either way, it is unlikely that language is extensively involved in our non-conscious reasoning, much of which is directed to the control of behaviour that is well within the scope of non-linguistic creatures. So we have a final division – between reasoning that is language-driven and reasoning that is not.

2.6 More links

Again, the tensions we have noted link up naturally into binary opposites: non-conscious reasoning may be non-explicit, can be characterized in Bayesian terms, is not intentional, and is rarely or never language-driven. Conscious reasoning, on the other hand, is explicit, is usually classical, can come under intentional control, and is often language-driven. Moreover, the two kinds of reasoning align nicely with the two strands of belief identified earlier: the former with the non-conscious, partial, passive, non-verbal strand, the latter with the conscious, flat-out, active, language-involving strand. Thus we can supplement our two-strand view of belief with a two-process view of reasoning, giving us a tentative two-strand theory of mind. For convenience, I shall refer to these two strands as *strand 1* and *strand 2* respectively.

3 MIND

In this part of the chapter I turn to two further divisions in the folk view of the mind, not so obvious at an everyday level, but soon apparent on philosophical reflection. The first concerns the ontological status of

Mind and Supermind

mental states, the second, the nature of mental explanation. Again, there is pressure to close these divisions by adopting a unitary approach, and again I shall suggest that a two-strand approach is preferable.

3.1 Ontology

Belief possession is typically associated with the possession of various behavioural dispositions. If you believe that something is true, then you will typically be disposed to act in ways that would be rational on the assumption of its truth. (What ways these are will, of course, depend on the circumstances and your background beliefs and desires; the dispositions associated with a belief are, in Ryle's phrase, *multi-track*.) Now, some theorists take belief ascriptions to refer simply to these multi-track behavioural dispositions: to ascribe a belief to a person is, they claim, first and foremost to say something about what the person is disposed to do. I shall refer to views of this kind as *dispositionalist*. (A more familiar term is *behaviourist*, but the word has unfortunate associations.) Other theorists, by contrast, take belief ascriptions to refer to the causal bases of these dispositions – to the states which *give rise* to the behaviour associated with the belief. I shall refer to these as *categorical-state* theorists. The most popular and attractive categorical-state theories – the various brands of functionalism – identify these states with functionally defined states of the brain.¹⁷

On the face of it, dispositionalist and categorical-state theories appear strongly opposed. They treat beliefs as very different kinds of thing – one as powers or tendencies of the organism as a whole, the other as states of its central nervous system. This opposition may be specious, however. For a strong case can be made for thinking of dispositional states and properties as functional ones. According to this view, to ascribe a disposition to an object is to ascribe to it a state or property with a certain causal role – a state or property which, in the right circumstances, causes the events which manifest the disposition. So in ascribing *fragility* to a glass we are ascribing to it a state which in the right circumstances causes shattering. If we couple this view with the idea that functional states are token-identical with the states that realize them (not *type*-identical, of course, since functions are multiply realizable), then we can regard token dispositions as identical

¹⁷ Among the dispositionalists, I count Davidson, Dennett, and Ryle; among the categorical-state theorists Armstrong, Fodor, and Smart. Note that throughout I assume a broadly *realist* view of dispositions. The case for this view is very strong (see Armstrong 1968, ch. 6; Mellor 1974; Mumford 1998, ch. 3).

Divisions in folk psychology

with their token categorical bases. On this view, then, the gap between dispositionalist and categorical-state views of belief narrows. Both sides can agree that beliefs are to be typed by their functional role, and both can identify token beliefs with token brain states.¹⁸

For all this, there is an important distinction lurking here, albeit not one best captured by the dispositional/categorical distinction. For though dispositionalists can be regarded as endorsing a form of functionalism, their functionalism is typically very different from that of categorical-state theorists. Dispositionalists think of beliefs as what I shall call *thickly carved* functional states – that is, as states of the whole cognitive system, defined primarily by their relations to inputs and outputs (perceptual stimuli and intentional actions). Thus, on a dispositionalist view, to possess a belief is to be in a state which produces a certain pattern of behavioural responses to perceptual stimuli, the nature of the responses varying with one's background beliefs and desires, themselves similarly characterized. However, the view involves no assumptions about the nature of this state or about the processes which produce the responses. It is thus compatible with dispositionalism that the internal basis of one belief may overlap in complex ways with those of others – indeed, that it may be nothing less than the whole cognitive system – and that reasoning may involve global patterns of neural activity. (Note that since dispositionalists must refer to the agent's background beliefs and desires in characterizing the role of a given belief, it follows that their definitions of mental-state terms will be holistically intertwined. This is not a problem, however: holistic intertwining of this kind is evident in many conceptual schemes; see Carruthers 1986, pp. 104–7.)

Categorical-state theorists, on the other hand, tend to think of beliefs as *finely carved* functional states. They regard them as functional sub-states of the cognitive system, defined not only by their relations to behavioural

¹⁸ Functionalist theories of dispositions are defended in Mumford 1998; Prior 1985; and Prior, Pargetter, and Jackson 1982. The account in the text draws heavily upon Mumford. Note that on a functionalist view the dispositional/categorical distinction is naturally understood as a *relative* one. A functional system may be realized in more basic, lower-level, functional systems, which are themselves realized in still more basic functional systems – and so on, all the way down (so-called homuncular functionalism or homunctionalism; see Dennett 1975; Lycan 1990). On this view, then, one and the same state or property may count as categorical relative to a higher level of organization, and as dispositional relative to a lower-level one (Mumford 1998, ch. 9). Whether dispositional properties must ultimately bottom out in genuinely categorical, non-dispositional ones is a matter of some controversy. For the view that they need not, see Blackburn 1990; Mellor 1974; Mumford 1998, ch. 10; and Popper 1957.

Mind and Supermind

outputs and perceptual inputs, but also by their *internal* relations to each other and to other mental states. On this view, action is the product of explicit reasoning involving occurrent beliefs and desires, and mental states are defined in part by their role in this reasoning. So to say that a person has a belief is to say that they have an internal state which can be activated in occurrent form and which then interacts in characteristic ways with other occurrent beliefs and desires, leading ultimately to action. Belief ascriptions thus carry implications about the structure of the cognitive system and the processes that generate action. (It is true that dispositionalists will also need to refer to other mental states in characterizing the role of a given belief, but they will view them quite differently – as background conditions, rather than as discrete causally interacting entities.)

As I said, the contrast here is not best captured by talk of dispositions and categorical states; it is better to think of it as a contrast between different varieties of functionalism. To characterize it, I shall speak of *austere* and *rich* versions of functionalism – or, more simply, of *austere* and *rich* views of the mind.¹⁹ I want to stress that I shall treat both of these positions as broadly *realist* ones. It is true that austere theorists are sometimes described as *anti-realists* about the mind – as holding that folk psychology is merely an interpretative device, rather than an empirical theory with ontological commitments (see, for example, Botterill and Carruthers 1999, ch. 2). Some austere theorists have encouraged this by describing their position as *instrumentalist*. (Dennett adopted this tag in some of his earlier writings, though he has since abandoned it and now stresses his realist credentials; see his 1991c.) However, I think that these descriptions are unhelpful – at least, given the way I have characterized the austere position. The disagreement between austere and rich theorists is not over the *reality* of mental states, but over their *nature*. The former regard them as multi-track behavioural dispositions, the latter as functional sub-states of the cognitive system (that is, as thickly carved functional states and finely carved ones, respectively). Of course, there is a sense in which austere theorists are anti-realists about beliefs; they deny the existence of beliefs *as conceived of by rich theorists* (or at least they do if they also subscribe to the unity of belief assumption). But to assume that this makes them anti-realists tout court is to assume

¹⁹ This terminology is an adaptation of that used in Horgan and Graham 1990. Note that ‘austere’ and ‘rich’ are broad terms, which cover a variety of more specific positions. Indeed, richness can be regarded as a matter of degree: the more complex one’s view of the internal functional structure of the mind, the richer it is. In chapters 6 and 7 we shall look at arguments for the view that folk psychology is committed to an even richer view than that described in the text.

Divisions in folk psychology

that the rich conception of belief is the only viable one. If we do not make that assumption, then we can accept that both groups are realists in their own way.

Austere and rich views are, on the face of it, straightforwardly incompatible, and they yield different accounts of who and what counts as a genuine believer. When used in an austere sense, 'belief' will have a wider extension than when used in a rich sense. Some systems which qualify as believers in the former sense will lack the right internal architecture to qualify in the latter. The two views also have different consequences for the epistemology of mind. On an austere view, all the factors relevant to determining an individual's belief state are overt. Beliefs are multi-track behavioural dispositions, and if a person's behaviour can be reliably and comprehensively interpreted as a rational expression of a particular belief, then they count as possessing it (Davidson 1975; Dennett 1981b). (If two or more equally good interpretations are possible, then we may have to say that the agent's mental state is simply indeterminate. This is not an unacceptable consequence, however (see Dennett 1991c), and in practice the scope for divergent interpretation will be limited, especially when long and complex behavioural sequences are considered.) On a rich view, by contrast, it is not the case that all the evidence for a person's belief state is overt. Beliefs are internal states which play a certain role in the processes leading to behaviour, and behavioural interpretation provides at best good but defeasible evidence for their existence.

Austere and rich views also have different implications for the possibility of irrationality. On an austere view, attributions of mental states carry a presumption of rationality. We regard a person as possessing a certain belief only if their behaviour can be interpreted as a rational expression of it. Likewise, we regard a behavioural episode as an intentional action only if it can be interpreted as a rational manifestation of the agent's mental states. Behaviour that cannot be interpreted in this way will have to be written off as non-intentional – as behavioural 'noise'. (There may be room for some flexibility here, given the vagueness in the notion of rationality, but stark or systematic irrationality is ruled out; see Dennett 1982.) On a rich view, by contrast, the link between mental states and actions is less tight, and there is scope for blatant irrationality. Beliefs and desires influence action by way of explicit reasoning processes, and these processes may occasionally go astray or be distorted by emotional or other factors. When this happens the mental states involved will lead to actions which they do not justify and which are thus genuinely irrational.

Mind and Supermind

The conflict between austere and rich views of the mind is well established in the philosophical literature – Dennett being the leading advocate of the former and Fodor of the latter (see, for example, Dennett 1987; Fodor 1987). However, the roots of the conflict are clearly detectable in folk psychology itself. We have already seen that the folk are at least partially committed to a rich view – that they hold that some beliefs require occurrent activation and influence action by way of explicit reasoning. (Further evidence for a folk commitment to rich functionalism will be reviewed in chapter 6.) Yet, as we have also seen, this commitment is not unqualified. Think again of the beliefs involved in non-conscious reasoning, such as those that guide routine driving behaviour or the actions of an expert chess player. The folk, I suggested, are not committed to the view that such beliefs require, or typically receive, occurrent activation. In these cases the folk appear happy to take a much more austere view. The tension is also evident in the way we use folk psychology. On the one hand, we ascribe folk-psychological states very liberally – not only to other humans, but also to animals, artefacts, and even plants (Dennett 1981b). This is often an extremely effective way of understanding and predicting their behaviour, and it is hard to deny its appropriateness. Yet, at the same time, we often feel inclined to say that animals and artefacts do not *really* possess the ascribed states – at least, not in the same way that we ourselves do. This strongly suggests that folk psychology has a dual function.

So here we have another tension in folk psychology. It is not particularly troubling in everyday situations, but it is not something a developed science of the mind could tolerate. It would be disastrous to permit a systematic ambiguity in one of its central concepts. So some revision or regularization of folk usage seems to be in order. Now, once again, the traditional approach here is to seek a unitary solution: to advance a global defence of either austere or rich functionalism. But another option would be to align the two views with the two strands of mentality identified earlier – to see the rich view as characterizing conscious thought, and the austere one its non-conscious counterpart. I shall return to this suggestion in a moment, after discussing a related matter.

3.2 Explanation

Questions about the nature of mental states are bound up with questions about the function of mental explanation. It is a commonplace that there are two ways in which citing an agent's mental states can explain their

Divisions in folk psychology

behaviour. It can do so either by rendering it *intelligible* – by showing how it is rational in the light of the agent's beliefs and desires – or by identifying its *cause* – by picking out some state or event which was causally responsible for it. It is uncontroversial that mental explanations have the former function. And it is plausible to regard them as having the latter, too. In trying to understand the causes of a road accident, for example, we might find it natural to refer to the beliefs of the drivers involved. And *philosophical* objections to a causal reading are now widely agreed to have been flawed (Audi 1973, 1985; Davidson 1963; Dretske 1989; Goldman 1970). Despite this, however, there remains scope for dispute about the *kind* of causal explanation that folk psychology offers – and the dispute corresponds closely to that between the austere and rich versions of functionalism identified in the last section.

I begin by distinguishing two kinds of causal explanation. I shall assume that causation is primarily a relation between events (understanding 'event' in the everyday sense as an episode involving change of some kind), and that the most basic kind of causal explanation is one that identifies a causal event. So, for example, we might explain a car's skidding by saying that it was caused by the driver's braking suddenly. I shall speak of such causal events as *dynamic causes*. But standing states can also be cited in causal explanations. For example, we might cite a car's *having worn tyres* in explanation of its skidding when the driver braked suddenly. This state (unlike many other states of the car) was causally relevant to the skidding. How to analyse claims of causal relevance is not completely clear, but I take it that they are closely bound up with counterfactual claims: if the car had not had worn tyres, then it would not (other things being equal) have skidded on braking. The presence of the state was a necessary condition for the causal event to produce its effect.²⁰ I shall refer to causally relevant states as *sustaining causes* of events.²¹

Return now to mental explanation. Beliefs are causes of action, but are they sustaining causes or dynamic ones? The answer depends on whether our view of the mind is austere or rich. The austere functionalist will say that beliefs are only sustaining causes. On an austere view, to have a belief is to have a disposition to produce certain behavioural outputs in response to

²⁰ For a counterfactual analysis of causal relevance, see LePore and Loewer 1987 and 1989; for an alternative approach, see Braun 1995.

²¹ I borrow the terminology of dynamic and sustaining causes from Audi (Audi 1993). Nothing hinges on its use, however. If you balk at speaking of standing states as causes, then think of them simply as causally relevant states.

Mind and Supermind

certain stimuli – this disposition being construed as a functional state. Now, we can regard these states as causal; each token disposition will be identical with some token realizing state, which causes the events that manifest the disposition (see Dennett 1981a, pp. 49–50). (The resulting explanations will, it is true, be fairly uninformative. To say that I stopped at the red light because I believed that a red light is an instruction to stop would be to say that I stopped because I was in a state which typically causes stopping at red lights. But not all causal explanations are informative, and if causation is an extensional relation, then it will be possible to redescribe the causal relation in a more informative way.) However, persisting states like this can be only sustaining causes, not dynamic ones. It is true that (if event causation is assumed to be primary) a sustaining cause will not manifest itself without some triggering event which serves as a dynamic cause of the ensuing effect, but austere functionalists cannot hold that belief–desire explanations refer to such events. On their view, beliefs and desires are just not the right sort of things to be triggering events. This is not to say that folk explanations of action will never pick out dynamic causes; those which refer to the impacts of external stimuli may do so. But *belief–desire* explanations will not. For example, suppose that I see a dog and run away, and that I do so because I believe that dogs are dangerous and want to avoid danger. Here, on an austere view, my seeing the dog will count as a dynamic cause of my action, but my beliefs and desires about dogs and danger only as sustaining causes of it. Of course, various internal events will occur in the mediating process between the perception and the ensuing action, and these might also be identified as dynamic causes of the latter; but austere functionalists cannot maintain that folk–psychological explanations serve to pick them out. On their view, the folk vocabulary simply does not carve things finely enough.²²

Rich functionalists, by contrast, are not so restricted. For they hold that the folk–psychological vocabulary picks out internal events as well as persisting states – occurrent beliefs and desires as well as standing–state ones. And these internal events can serve as dynamic causes of action.

²² Thus Davidson, defending a broadly austere position, explains that the events which cause actions are those which *initiate* the intentional states that rationalize them:

In many cases it is not difficult at all to find events very closely associated with the primary reason [for an action]. States and dispositions are not events, but the onslaught of a state or disposition is. A desire to hurt your feelings may spring up at the moment you anger me; I may start wanting to eat a melon just when I see one; and beliefs may begin at the moment we notice, perceive, learn, or remember something. (Davidson 1980, p. 12)

Divisions in folk psychology

Indeed, the contrast with the austere position is even more marked. For on a rich view, not only *can* belief–desire explanations advert to dynamic causes; they typically *will* do so. This is a corollary of how rich functionalists conceive of beliefs and desires. On their view, to have a standing–state belief with content *p* is to be in a persisting state such that, given appropriate stimuli, either external or internal, one will entertain an occurrent belief with content *p* – an event which may then be the dynamic cause of further occurrent thoughts or overt actions. That is to say, on their view, standing–state beliefs are not dispositions to act, but dispositions to have occurrent thoughts, which may then in turn cause actions. So standing–state beliefs will be sustaining causes of occurrent thoughts, but not (or not directly) of the overt actions which those thoughts cause. Similarly with desires. So, if the aim of belief–desire explanation is to identify the causes of actions, rather than the causes of those causes, then successful explanations of this kind will refer to occurrent beliefs and desires, and therefore to dynamic causes.

The picture here is complicated slightly by the fact that a dynamic cause may be effective only in the presence of certain background conditions, which will count as sustaining causes of the resulting effect. And in the case of occurrent beliefs and desires these background conditions may include the presence of various beliefs and desires that are not occurrently activated (the suppressed premises and background assumptions mentioned earlier). So, for example, a background condition for my desire for bread to cause me to set off for the shop might be that I have the belief that it is safe to go out, and this belief could thus be cited as a sustaining cause of my action. So it is not true to say that on a rich view of the mind, belief–desire explanations will *always* advert to dynamic rather than sustaining causes. Still, they *typically* will. As I noted earlier, suppressed premises and background assumptions are things we take for granted – things we treat as part of the normal background to an episode of reasoning. And it will rarely be informative to cite such beliefs in explanation of an action. (Though there will be exceptions; suppose, for example, that you were explaining my shopping trip to someone from a war-torn country, where every journey out was fraught with danger.)

So austere and rich views of the mind each support a different sort of psychological explanation. It follows that if, as I have claimed, the folk shift between the two views, then we should expect to find them shifting between the two kinds of psychological explanation, too – sometimes aiming to pick out dynamic causes, sometimes content to identify sustaining

Mind and Supermind

ones. And vice versa: if we do find this duality in folk explanatory practice, then this will tend to confirm that the folk alternate between austere and rich views of the mind.

And this is indeed what we find. Some intentional explanations clearly aim to identify dynamic causes. Someone watching me drive to work and noting my sudden deviation from my normal route might seek to know what event had precipitated the change. And this inquiry might naturally be answered by saying that it had suddenly occurred to me that roadworks were due to start – an explanation adverting to a conscious occurrent belief which triggered the action. Other cases are different, however. Suppose that a novice watching a championship chess match wants to know why the players make the moves they do and why they ignore others which seem on the face of it more attractive. An expert replies by providing them with information about the players' mental states – their strategic aims, both short-term and long-term, their beliefs about the rules of the game, their knowledge of the standard openings and strategies, and so on. Here it is less plausible to think that the aim of the explanation is to identify dynamic causes. The expert is not claiming that the players occurrently entertained the beliefs and desires referred to – certainly not as *conscious* occurrent thoughts, or even, I suggest, as non-conscious ones. At any rate, the explanation would not be vitiated if it could be shown that the players had not entertained such thoughts. For the questioner is not much interested in the sequence of causal events that led to the players' actions – which will in all probability have been exceedingly complex. Rather, their concern is to understand the *rationality* of their actions – to see why their moves were wise ones for them to make. Of course, it matters that the explanation given be *true*; the questioner does not want just any old rationale for the players' behaviour, but the actual one. It is sufficient for that, however, that the explanation picks out sustaining causes of the players' actions – that the players possessed the beliefs and desires cited and would not have made the moves they did if they had not.²³

3.3 From theory to theories

It appears, then, that belief–desire explanation has a dual function – sometimes picking out dynamic causes, sometimes sustaining ones. This in turn

²³ For further evidence that folk-psychological explanation does not require a strong causal reading, see Anscombe 1957, sect. 11.

Divisions in folk psychology

confirms the earlier suggestion that folk mental concepts have both austere and rich versions. The upshot is that we can separate out two different folk theories – one austere, the other rich – each supporting a different conception of the mind and mental explanation. Now, this would not in itself compel us to adopt a two-strand theory of mind. We might see the two theories as describing different *aspects* of a single cognitive system. We could regard one as a *competence theory*, which characterizes the powers of the system at a shallow level, and the other as a *performance theory*, designed to identify the causal mechanisms supporting those powers.²⁴ The two theories would differ in scope and falsification conditions, and their concepts would have different extensions, but each might be refined and regularized to meet the standards appropriate for theories of its type.

From our present perspective, however, there is another – and, I think, more attractive – option. For the distinction between the two versions of folk psychology corresponds at least roughly to that between the two types of belief identified earlier. The evidence for a rich view stems mainly from conscious thought, that for an austere one from its non-conscious counterpart. So another way to regularize folk practice would be to think of each theory as characterizing one of the two strands of belief we identified earlier – that is, to adopt an austere view of strand 1 belief and a rich view of strand 2 belief. I propose, then, that a revised folk psychology should incorporate two sub-theories: an austere theory of the strand 1 mind, which picks out thickly carved functional states and sustaining causes of action, and a rich theory of the strand 2 mind, which picks out finely carved functional states and dynamic causes. Of course, this involves rejecting the idea that the sub-theories, austere and rich, are related as competence and performance theories – that the latter characterizes the causal mechanisms supporting the states and processes described by the former. It would be absurd to suggest that non-conscious states and processes are supported by conscious ones! Rather, we must think of the two theories as characterizing distinct strands of mentality.

Let me stress that this two-strand framework is offered as a regularization of *folk* psychology, not as an exhaustive psychological taxonomy. It

²⁴ The terminology of competence and performance derives originally from Chomsky, who employs it to mark a distinction in linguistic theory. I use the terms here in the looser sense adopted by Daniel Dennett (Dennett 1981a). Dennett illustrates the use of the terms by reference to the physicist's distinction between *kinematics* and *dynamics* – the former providing an idealized, abstract level of description, and the latter a deeper, genuinely causal one.

Mind and Supermind

may be that there are other strands of cognition, unfamiliar to the folk, and that we shall need new branches of intentional psychology to describe them. Indeed, there will almost certainly be a level of *sub-personal cognitive psychology* underlying the strand 1 mind. (Here the terminology of competence and performance really is appropriate. Theories of sub-personal psychology will aim to characterize the causal underpinnings of the dispositions picked out by the austere folk theory.) Moreover, the sub-personal system might turn out to share some properties with the *strand 2* mind. It might consist of discrete representational states which can be occurrently activated and which enter into explicit inferential processes. It might even exploit the representational resources of the language system. Nothing I have said here rules that out. My claim is merely that *folk psychology* does not involve any commitments as to the nature of this level (except negative ones – that it is not conscious, not under active control, and so on). The folk are not concerned with the character of sub-personal cognition, and it is precisely their nonchalance about it that makes an austere view of the non-conscious mind so attractive. (This is not to say that people never speculate about sub-personal processes – just that the core practices of folk-psychological explanation and predication do not depend on assumptions about them.)²⁵

Does this mean that we should regard the folk theory of the strand 1 mind merely as a placeholder for a scientific theory of sub-personal cognition? In a sense, yes. If we want to understand the workings of the non-conscious mind, then we shall have to move beyond the austere folk perspective. But for many purposes that perspective may be perfectly adequate. For the high-level sciences of human behaviour, such as sociology and economics, the details of sub-personal cognition will often be largely irrelevant. And for everyday use, too, it is likely that theories of sub-personal cognition will be unwieldy and redundant, though elements of them may be incorporated into folk use, especially if they shed light on common pathological conditions. (This happened with Freudian psychoanalytic theory, and I suspect that it will happen with modern evolutionary psychology.)

Finally, note that the proposal just made puts the earlier discussion of strand 1 belief in a slightly different light. I suggested that there is no pre-theoretical reason to suppose that strand 1 beliefs receive occurrent activation or involve natural language, though I did not rule out the possibility

²⁵ For discussion of the relation between folk psychology and sub-personal cognitive psychology, see Dennett 1981a.

Divisions in folk psychology

that they might. However, if these states are simply behavioural dispositions, then we can indeed rule this out. Behavioural dispositions manifest themselves directly in action, not in episodes of occurrent thought. Nor does it make sense to think of dispositional states as employing a linguistic medium – such a claim can be made only for occurrent thoughts or stored representations. It remains possible, of course, that some of the sub-personal states and processes underlying the strand 1 mind are explicit and language-involving.

3.4 A Bayesian mind?

With our two-strand theory fleshed out, I want now to return briefly to a couple of issues postponed from earlier, both relating to the strand 1 mind and the role of Bayesian idioms in characterizing it.

The first issue concerns the nature of partial beliefs. I promised to say a little more about what these are, and, having argued for an austere view of the strand 1 mind, I am now in a position to do this. Partial beliefs are Bayesian subjective probabilities, and these states in turn are multi-track behavioural dispositions, understood as thickly carved functional states (functions from inputs to outputs). To have a certain set of subjective probabilities is to be disposed to make the choices that a rational Bayesian agent with those probabilities would make, given one's subjective desirabilities. Since one can have such a disposition without having made explicit judgements of probability, it follows that no special conceptual apparatus is needed in order to possess partial beliefs. Likewise, partial desires are subjective desirabilities, understood in the same way.

The second issue concerns strand 1 reasoning. I claimed that this could be characterized as Bayesian, but noted an objection to this claim. Just because we can be *represented* as Bayesian reasoners, the objection ran, it does not follow that we *are* Bayesian reasoners – that Bayesian theory describes real psychological processes. Given an austere view of the strand 1 mind, however, this objection loses its force. For on this view a theory of inference serves as a framework for behavioural interpretation rather than as a model of internal processes. We credit people with the mental states which provide the best overall interpretation of the choices they make, on the assumption that these choices are related to their mental states in the way prescribed by the theory, under its normative aspect. There is, however, no assumption that their choices are the product of actual calculations of the sort specified in the theory: the theory characterizes

Mind and Supermind

the agent's behavioural dispositions, but not the processes which produce their behaviour. From this perspective, then, a theory of inference has only a shallow descriptive role, and being representable as a Bayesian reasoner is sufficient for being one.

The reader may ask why, given this general approach, a Bayesian interpretation should be privileged. There will always be alternative ways of interpreting an agent's choices, which represent them as having different attitudes and as adhering to different norms (for a demonstration of this, see Zynda 2000). So what grounds are there for regarding the Bayesian representation as giving a privileged description of reality? Indeed, why not interpret the agent as having *flat-out* beliefs and desires and adhering to the norms of classical practical reasoning? In short, how do we choose which normative theory to use as the basis for interpretation? The right response here is to say that the choice of interpretative framework is a pragmatic matter. If rival frameworks differ in their behavioural predictions, then we can choose between them on the basis of their success. If they are predictively equivalent, then we can treat them as alternative ways of characterizing the same underlying dispositions, and choose between them on grounds of simplicity and conformity to our pre-theoretical intuitions.²⁶ I strongly suspect that these criteria will dictate a probabilistic framework of some kind, rather than a classical one. It seems likely that a classical framework would be severely limited in its descriptive and predictive power unless the attributed beliefs and desires were assigned degrees of strength. And from an austere perspective, it is unclear how a qualified classical framework of this kind would differ from a probabilistic one.²⁷

It may be objected here that there is evidence that human reasoning cannot be interpreted as Bayesian. There is a large experimental literature in the 'heuristics and biases' tradition showing that in certain test situations people regularly respond in ways that violate Bayesian principles (see, for example, Nisbett and Ross 1980; Kahneman et al. 1982; Piattelli-Palmarini 1994). Moreover, the responses of the test subjects seem to be the product of intuition rather than conscious reasoning – suggesting that Bayesianism

²⁶ Note that to say that the choice of interpretative frameworks is a pragmatic matter is not to say that psychological descriptions are observer-relative. As Dennett points out, different interpretations can each highlight different, but equally real, patterns in an agent's behaviour (that is, distinct, though compatible, behavioural dispositions) (Dennett 1991c). Dennett is thinking of different interpretations *within* a particular normative framework, but the point applies equally to ones that rely on different frameworks.

²⁷ Davidson also argues that behavioural interpretation requires a probabilistic framework; Davidson 1975.

Divisions in folk psychology

is an inappropriate interpretative framework for the non-conscious strand of mentality. (The results *also* suggest that the underlying sub-personal processes are not Bayesian. But that claim is not in itself troublesome, for the reasons given above.) There are two broad ways of responding to this. The first is to stick with a Bayesian approach and write off any recalcitrant behaviour as ‘noise’. What is required for interpretability is a broad adherence to norms, not an exceptionless one. After all, there will always be some recalcitrant data, whatever interpretative framework we employ. (Moreover, it may be that the heuristics-and-biases results are less embarrassing than they seem at first sight. There is evidence that the subjects’ errors are due to the artificial format of the test problems, and that when the same problems are posed in a more familiar form, the responses are much more in line with Bayesian principles: see Gigerenzer 1991; Cosmides and Tooby 1996.) The second response is to employ an interpretative framework which is tailored to reflect the capacities and functions of our sub-personal cognitive mechanisms and which is thus a more accurate predictor of actual human responses (see Stich 1982). This framework would still need to be a probabilistic one, for the reasons given above, but it might differ in significant ways from standard Bayesian theory. I suspect that a strong case can be made for this second option, but for simplicity’s sake I am going to adopt the first and take Bayesianism as the default position. However, nothing in what follows depends on a commitment to strict Bayesianism, and another normative framework could be substituted for it without substantially affecting the arguments.

There is another issue I want to address briefly. Since I am proposing a two-strand theory of mind as a regularization and development of folk-psychological practice, it is important that folk idioms can be retained in characterizing both of the strands. But on the proposed Bayesian framework, they appear ill-equipped to characterize the strand 1 mind. The problem concerns mental explanation. The folk practice is to single out a small number of beliefs and desires in explanation of an action. But if we interpret strand 1 actions as the product of Bayesian reasoning, then it is not clear that this practice can be sustained. For Bayesian reasoning is holistic – the product of *all* our partial beliefs and desires. What role is left for the folk practice of singling out individual ones?

There is, I think, a perfectly good role for it. The crucial point to remember is that, on an austere view, mental explanation singles out *sustaining* causes of action, not dynamic ones, where claims about sustaining

Mind and Supermind

causation can be cashed out in counterfactual terms. Thus, when we cite a partial belief or desire in explanation of an action, we are saying that the action was counterfactually dependent on its presence. And, given this, the folk practice makes good sense. For an action may be counterfactually sensitive to changes in some, but not all, of the agent's partial beliefs and desires. Take the case where I see a dog and run away. And suppose that I would not have run away if I had attached a much lower probability to the proposition that dogs are dangerous, but that I would still have run away if I had attached a much lower probability to the proposition that Tallahassee is the state capital of Florida. Then this fact alone justifies us in singling out my attitude to the former proposition rather than my attitude to the latter in explanation of my action – in saying that I ran away because I believed that dogs were dangerous. (Note that the claim here is not that my action would have been sensitive to *any* change in my attitude to that proposition – only that it would have been affected by *substantial* ones. If I had been just a bit less convinced that dogs were dangerous, I might still have decided to run.)

It may be objected that if strand 1 reasoning can be interpreted as Bayesian, then actions will be counterfactually sensitive in this way to changes in *any* of the agent's partial beliefs and desires. For Bayesian reasoning is holistic – its outcome determined by all of the agent's probabilities and desirabilities. This is too swift, however. We need to distinguish two different outcomes of the notional Bayesian process: the complete assignment of estimated desirabilities to the candidate actions, and the determination of which action has the *highest* estimated desirability – which comes top in the estimated desirability ranking. If the process is holistic, then the complete assignment of estimated desirabilities will indeed be sensitive to changes in any of the agent's probabilities and desirabilities: if any of them had been different, then the assignment would have differed in some way. But it is very unlikely that the determination of which action comes top will also be sensitive to any such change. In my dog encounter, for example, it seems likely that *running away* would still have come out as the most attractive action, no matter what probability I had assigned to the proposition that Tallahassee is the state capital of Florida. So an action's coming top, and therefore being executed, will be counterfactually sensitive to substantial changes in some, but not all, of the agent's probabilities and desirabilities, and we may legitimately single out these states in explanation of it. It is true that an action will be counterfactually sensitive in this way to *many* probabilities and desirabilities, few of which we would

Divisions in folk psychology

normally think of citing in explanation of it. For example, my decision to run away from the dog would be sensitive to substantial changes in the probability I attach to the proposition *I can run*. But this is by no means an unacceptable consequence; explanation invariably involves picking out a few of the more salient factors from among a host of others.

I conclude that the adoption of a Bayesian view of the strand 1 mind is not incompatible with the retention of folk-psychological explanatory practices – though it does require them to be reconstrued. In the following chapters I shall from time to time follow the folk practice of singling out individual beliefs and desires in explanation of actions generated at the strand 1 level – always on the understanding that such idioms are to be construed in the way just described.

I shall address one final worry. In adopting an austere view of the strand 1 mind, I am committed to anti-realism about strand 1 mental processes. Strand 1 states may be real dispositions, but strand 1 processes are merely notional – an interpretative fiction, employed to characterize strand 1 states. And this may seem implausible. Surely, when we talk of non-conscious mental processes we mean to refer to real processes, not notional ones? (Indeed, a commitment to realism was implicit in my own earlier discussion of strand 1 reasoning – in the discussion of whether it was explicit or non-explicit, active or passive, and so on.) I concede the point: there is a commitment to realism here. However, it can be reconciled with the austere perspective advocated. When we refer to non-conscious mental processes, we are, I suggest, referring to the real sub-personal processes *in virtue of which* we are interpretable as Bayesian reasoners. We take no stand as to the *nature* of these processes, however, but simply quantify over them (or if we do characterize them, do so in negative terms – as not under active control, not involving inner speech, and so on). In this respect, the view remains perfectly austere. Henceforth, references to strand 1 reasoning or to non-conscious mental processes should be understood in this way.

CONCLUSION AND PROSPECT

I have provisionally identified two strands of belief, associated with two kinds of mental processing and two conceptions of mind and mental explanation. Their characteristics are summarized in figure 1.

Let me emphasize that I am not offering this neat and tidy formulation simply as an *analysis* of folk-psychological practice, but as a *theoretical*

Mind and Supermind

	Strand 1	Strand 2
Belief	<ul style="list-style-type: none"> • Non-conscious • Not apt to be activated in occurrent form • Partial • Passively formed • Not language-involving • Common to humans and animals 	<ul style="list-style-type: none"> • Conscious • Apt to be activated in occurrent form • Flat-out • Can be actively formed • Frequently language-involving • Unique to humans and other language users
Reasoning	<ul style="list-style-type: none"> • Non-conscious • Interpretable as Bayesian • Depends on sub-personal processes that are not under active control, may be non-explicit, and are probably not language-driven 	<ul style="list-style-type: none"> • Conscious • Usually classical • Can be actively controlled • Explicit • Frequently language-driven
Mind	<ul style="list-style-type: none"> • Mental states are thickly carved functional states (austere functionalism) • Belief–desire explanations pick out sustaining causes 	<ul style="list-style-type: none"> • Mental states are finely carved functional states (rich functionalism) • Belief–desire explanations typically pick out dynamic causes

Figure 1 The two strands of folk psychology

regularization of it. The roots of it are present in folk discourse – the divisions are there, as are the links between them. And this way of ordering things is, I think, the natural one. But I do not doubt that counter-examples could be produced – instances of folk locutions which suggest a different classification, conjoining properties and processes which are here labelled distinct. I am not offering this as the *only* classification possible, but as the best and most consistent one.

We have a framework, then, which nicely organizes some of our common-sense intuitions about the mind. But it raises as many questions as it answers. First, there are constitutive questions. I have suggested that strand 1 beliefs are realized in sub-personal states and processes, but what of the strand 2 kind? I have claimed that they are functional sub-states of the cognitive system, but how are they realized? It is widely held that beliefs are token-identical with brain states, but this is not essential to the folk outlook. (It is quite coherent to claim that beliefs are states of a non-physical soul.) And even if we assume (as I shall) that mental states are physically constituted, we do not have to regard them as realized directly in states of the brain. Secondly, there are questions about the relation between the two strands of mind. How are the two belief systems

Divisions in folk psychology

related to each other, and what are their respective roles in the guidance of action? Thirdly, there are questions about the *function* of the two strands. Why do we need two strands of cognition? If complex tasks such as driving and chess-playing can be controlled non-consciously, what is the role of conscious reasoning? In short, we need to convert this outline framework into a coherent and plausible psychological theory. This will be the task of the next three chapters.