The Problem of Induction and the Problem of Other Minds:

A Proposed Solution

PhD thesis 2011

Laurence Peddle

Summary

In chapter one I reject the thesis that a relation of partial entailment holds between probability evidence and conclusion, as also the claim that the propositions of applied geometry are synthetic *a priori*. By way of balance, I suggest that there may be a sense in which necessary conditions of discourse transcend the distinction between *a priori* and empirical.

In chapter two I reject the claim that logical relations in the form of intrinsic probability enter into the no-miracles argument, which I suggest is frequency-connected in its more systematic applications, so that it belongs within a system.

I begin chapter three with a critique of an attempted formal probability solution to Hume's problem, and I now suggest that inductive inference has application only within a system in which its validity is pre-supposed in its premises, a concomitant of which is that the sceptic about induction cannot stop short of global scepticism.

Since my aim is to show that global scepticism is self-refuting, given that intentionality may be analysed in terms of a system, I now develop that analysis by devoting chapter four to an examination of Wittgenstein on meaning and understanding.

In chapter five I reject his thesis equating meaning and understanding with use, arguing instead that they are irreducible and subject to dispositionality conditions, and in furtherance of that argument I try to solve the problem of the authoritativeness of belief avowals by showing again that there are necessary conditions of discourse. These are such that self-ascribing belief, crediting oneself with understanding and with being suitably disposed, are inherent in reasoning.

In chapter six I weave the threads of the previous discussion into a solution by arguing that the sceptic about induction, who is committed to global scepticism, necessarily refutes himself.

In chapter seven, on the problem of other minds, I attempt a solution by modifying the arguments used against inductive scepticism.

Contents   Page

<u>Acknowledgements</u>

My grateful thanks to my thesis supervisor, Professor Christopher Norris, for all his support and encouragement over so many years and in so many ways. These include his always having an open door whenever I wished to check on my progress or ask his advice. In every chapter of this work I have benefited from his detailed comments and helpful remarks; and, too, my thinking has been sharpened by the cut and thrust of debate in the reading group sessions he has led. I owe a great deal, also, to the discussions I have had at conferences, seminars and elsewhere with fellow students and others too numerous to mention by name.

I would, however, like to mention my wife's name and to thank Helen for all her sacrifice and support and for eventually accepting my claim to be working hard on my thesis even when I was not reading or writing and appeared to be neglecting my studies. She has come to understand the crucial role played by unconscious thinking, which my own research indicates to be at its most effective when study is combined with dancing, tennis and gardening. She has been kind enough, also, to draw up a long list of all the do-it-yourself jobs which over the years both of us have noticed needed to be done, and which she has now listed in order of priority with detailed instructions and a work schedule.

Introduction

Since the title of this thesis announces a belief on my part that epistemological problems are soluble, I should perhaps begin with an account of what I hope to achieve. My declared aim is to refute the sceptic about induction and other minds, but to that end I need also to deliver our knowledge of the past from him, solve a puzzle about the authority with which we self-ascribe belief and other intentional states, and shed light on the epistemic status of experiential statements and observation reports. This programme can be carried through only if certain distinctions, between induction, other forms of factual inference, and physical description, are divested of the significance normally attributed to them; also, my over-arching aim is to exhibit our reasoning about the world as conducted within a system constrained by necessary conditions of discourse. In pursuing these aims, I hope to be able to discuss the various issues in the same arena as that occupied by contemporary debate on them, and thereby to contribute to it.

In my critique of Wittgenstein on intentionality I take issue with his professed aim of dispelling misconceptions about meaning and understanding, for it seems to me that the proper task of philosophy in this arena is that of enquiry and analysis. Wittgenstein, however, seems to exhibit empiricist leanings in his private language argument and in his general approach to what it is to mean and to understand. In equating meaning with use, he maintains that semantic content does not consist in the mental phenomena associated with giving meaning to a sentence, a point of view which on one interpretation is tantamount to semantic non-factualism. This is radical scepticism about meaning, and I argue that it should be rejected, its place taken by the thesis that semantic content is irreducible and sui generis: in other words, and in line with common sense, that there are facts about meaning and that they fall into a category of their own. By way of illustration, if here in the attic I smell smoke and hear a roaring noise downstairs, whereupon the sentence 'Oh no, there's a lion in the house and one or both are on fire!' comes into my mind at the same time as I think that there is a lion in the house and that one or both are on fire, then in normal circumstances[1] the sentence will express the thought, which will be occurrent, as will my awareness, non-verbal, of its having been occasioned by the smell and the roaring noise.

---

[1] Abnormal circumstances could be those in which, for instance, I have the thought but do not understand the sentence, perhaps because I do not speak English, so that I am not even aware of the fantastic coincidence.

The point to be made is that once the sceptic has been defeated, so that the concept of intentionality is no longer under siege, then we are free to explore it in a spirit of enquiry, one finding of which is that intention states combine the occurrent and the dispositional, which therefore *cannot* conflict, the task of analysis being to tease out their interconnections. This may be done by developing a theory to which that analysis contributes, and to that end I have employed the notion of a system, in terms of which the theory may be outlined as follows, the reference being to understanding rather than to meaning. First, we communicate with others only if we communicate with ourselves, so that communal understanding derives from that of the individual, which is therefore the basis from which we proceed. The elements to be married up are the occurrent, the dispositional and the irreducible, together with conditions of discourse. If, to argue from a particular case, I understand the even number sequence, then my understanding is occurrent, for I must now be suitably disposed, and if I now exercise that understanding, perhaps by writing down the first few terms of the sequence, then again it is occurrent. Since understanding is irreducible, it must inform my dispositions, so that I am disposed to manifest my understanding in a variety of ways, the same analysis being applicable to each actualisation. Thus, a condition of understanding the sequence is that one should be disposed to understand it, from which it follows that dispositions determine the content of that understanding. What also follows, from a first person perspective, is that both understanding and the lack of it obtain within a cognitive framework, so that a modicum of understanding is required even to think, correctly or not, that one understands or fails to understand the sequence. Thus, it is a necessary condition of discourse that one should think that one understands. There are other necessary conditions, the details of the theory being found in the main text.

This is, no doubt, a debatable theory, and in support of it I engage in argument with Kripke as well as Wittgenstein. Also, I critically examine the views of Hattiangadi, Wright and McDowell on such matters as semantic non-factualism, non-reductionism and normativity. My ultimate aim is to solve the induction problem, and in chapter 6, having already established that the inductive sceptic cannot stop short of global scepticism, I set out my attempt at a solution. Agreeing with Okasha and others that Hume has misunderstood the nature of factual inference, I maintain that nevertheless he can be refuted only if his arguments entail global scepticism. My approach is to show that conditions of meaning and understanding are such that the global sceptic is committed to claiming knowledge of the past and the future, so that he necessarily refutes himself. Having achieved this result by applying the notion of a system, I now turn it to further use in my treatment of the challenge posed by

sceptical hypotheses, such as that of the brain in a vat. I argue first of all that a contextualist response to the problem is unsatisfactory. Applying the notion of a system, I then try to show that the general fact exploited by the sceptic: that in all cases empirical knowledge is underdetermined by evidence, is fundamental to discourse about the physical world, a consequence of which is that the implications of any sceptical hypothesis are such as to render the sceptic's argument incoherent. In the closing pages of the chapter, I consider Wittgenstein's epistemology as presented in *On Certainty*, and I conclude that although his use of the notion of a system has affinities with mine, he is unable to demonstrate that inductive scepticism is self-refuting. Finally, I return to the analysis of factual inference and try to show that Goodman's arguments about deviant predicates do not reveal a genuine paradox about projectibility, and for the reason that empirical knowledge belongs within a conceptual and explanatory system in which the notion of linear extrapolation from repeated instances has no place, given that Hume's theory of the nature of inductive inference was misconceived.

This brings us to the chapter on the epistemological problem of other minds, and here I have examined Cassam's theory of epistemic perception constructed around the fact that when we observe others we 'see' them as having particular mental states, this being associated with our saying that we see that they are, for instance, in pain. Having rejected that theory as an answer to scepticism, along with his response to sceptical hypotheses targeted at knowledge of other minds, such as that other people are zombies, I then turn to my own attempt at a solution. Here I take the same approach as with the problem of induction, and use the notion of a system to counter foundationalist and empiricist assumptions about conceptual and epistemological difficulties with regard to our knowledge of others' mental states. There is no fundamental distinction, in my view, between knowledge of ourselves and of others, or between what we know of the past, the present and the future. This is implied in the theory that knowledge of the world forms a system, the fundamentals of which are found not in such distinctions but in the necessary conditions by which reasoning about the world is possible. These are such that the taking for granted of the deliverances of memory, expectation and perception are inescapable and beyond question, doubts being raised only in particular cases in which the alternatives considered all rely on these epistemic methods and sources.

Once the epistemological problems with which we have been concerned have been dealt with, our attention should, in my opinion, switch to the metaphysical implications of that notion of a system which has been instrumental in our treatment of those problems. In supplanting a world-view based on empiricist and foundationalist assumptions, it opens up a

wider vista in which interconnections replace foundational empirical data, a question then arising as to what is being connected and what drives it. We cannot say that the nodes, as if it were a network not a system, are perceptual objects, for to identify such an object is to go beyond what is given, since nothing is. When I glance repeatedly at a table, always expecting it to be there, such expectations are intrinsic to my identifying it as a table in the first place, since inference enters into perception and belief. All we know are effects, not causes, except that this is to use concepts belonging to the surface of a world in which deeper forces have surface effects, for we have no choice. All the same, we are free to try to give definition to what cannot be said in terms of what can, and to express ourselves in one way and then another.

To that end, let it be said that a fog of unknowingness envelopes the metaphysical implications of the concept of a system; and, if we cannot penetrate it then perhaps we can at least find the words with which to depict it, so that we are not completely at sea. If we jettison the assumption that intrinsic inferential forces bind regularities and repeated instances into firm anchorage from which to shine a light over the horizon, then all that is left is the sea anchor of empirical data, except that it, too, has unknown depths and is dragged along with us by unseen currents and cannot secure our beliefs. Patterns form and break on the surface of a perceptual ocean, but their hidden engine works far below, its experiential effects mediated by the senses, which themselves are its effects. Sometimes the fog lifts and swirls around vaguely discerned questions, but we cannot give enough shape to them to be able to formulate them, except to say, in the space where answers should be, that they concern the nature of this world and of the place we occupy in it. The best I can do, in such poor visibility, is to try to represent the nature of things as being fundamentally mysterious, and to suggest that this is to draw upon a universal sense of the incomprehensibility of it all which for practical purposes we keep in check – or translate, perhaps, into philosophical puzzlement.

In evoking a sense of mystery, then, perhaps we may begin with the observation that not even the most cogent refutation of the sceptic, let alone my own modest proposal, can succeed in permanently silencing him, partly because the philosopher is not a scientist, and although his discipline may change and evolve, this is not to say that it progresses in the sense of adding to an established body of fact. As for why this should be the case, we could appeal to the potential for conceptual conflict in each of us, with concepts working smoothly enough, perhaps, in their unexamined state and in their own sphere, but jarring against one another when coming under scrutiny. The habit of pushing against conceptual limits, or of

one concept asking questions of another, is no doubt nurtured by philosophising, but its roots go deep enough to tap into the minds of all those who contemplate the inscrutability of this world.

The central point, then, is that the surface intelligibility of the deliverances of perception arises from a deep space of things unknown, the presence of which we are more fully confronted with if we embrace the notion of a system, having been swept from the rock of factual knowledge being empirically based in some fundamental sense. To develop that point, let us begin with the fact that we apprehend physical entities as being spatio-temporally extended, this awareness being not just theoretical but intrinsic to processes of perception. We 'see' a physical object as being three-dimensional, which is also the way in which our sense of touch interprets it, so that we seem to have immediate acquaintance with it as an object in space and time; and yet, we also have a sense of physical reality as inhabiting only the present moment, except that it tantalises us by instantly being swept away. It is as if all that we can grasp, really, are the resemblances, as Hume would say, which link successive moments in a chain of awareness, except that this is to appeal illicitly to the notion of discrete units of time, for which we may try to correct by acknowledging that each link is broken in the act of being formed. Even so, if we speak of the flow of temporal reality, and then of it proceeding in linked steps, then such metaphors cannot be married up and are in any case inadequate, as any others would be, so that the mystery remains. One may speak of the specious present, but the mystery goes deeper than a change of terminology could resolve.

In fact, it may go deeper again, or the clash of perspectives may, if we now turn to our paradoxical view of the very idea of the past, the previous comments on the reality of the present moment also being relevant here. Let it be said first that we are never more at home with our sense of the physical reality of things than when examining a medium-sized solid object of stable appearance, preferably one that we can lift and handle, such as a pebble. A presently examined pebble, then, fits snugly against the ontological benchmark by which we judge the reality of the contents of this world, whereas those belonging to the past fall short in a way that cannot be remedied by change of tense. Thus, I say that yesterday for philosophical reasons I plucked this pebble from the beach, and I thereby refer to an event which *was* real, just as my now licking the pebble, having washed it first, *is* real; but my memory of yesterday's excursion to the seaside, and of this pebble found then, lacks all solidity and yet is vivid and alive, so that in limited ways the past is recreated in the present – except, of course, that it is not. This, surely, is the kind of dissonance that unsettles us all,

especially if the background theme is one of love or loss, as when we are haunted by our memories.

What we have, then, is that the past may seem unreal compared with the present, or a presently observed object or event, despite our grip on the felt reality of the present itself being uncertain if an undercurrent of doubt is at work, given that the present is instantly swept away into the past. And yet, our feel for the reality of what is presently observed, especially if it is a pebble rather than a rainbow, inclines us to the view that experience of the external world in which that reality inheres is a vital component of our conception of that world. So much so, indeed, that Hume speaks of ideas arising from impressions, and even the non-philosopher is predisposed to the view that items in the abstract realm, such as numbers and symbols, are deficient in the quality of being real – and yet, nothing could be more abstract than my understanding of this pebble, worn smooth by the wash of language over thousands of years of speech.

If the edifices of the empiricist, the realist and the sceptic, together with those of their rivals, arise from such discontinuities of thought, or from the deep structures of the world, then we may continue to dig down to them, to describe but not to understand, if we now consider the role of perception in constructing that world. Starting from the shallows of naive realism, and then with a modicum of immersion in philosophy, we quickly discover that we are not so buoyant in these deeper waters, the old assumptions being dragged under by new ideas, and that the objective reality of the external world can no longer be taken for granted. We have learnt that a straight stick looks bent when partly submerged, and that a mountain changes shape as we circumambulate it, or shrinks amazingly to nothing when we close our eyes or turn away. Still, at least there is *something* that we see when we look again, except that it may now occur to us, though it is often overlooked, that if we query whether the physical world is real, our acquaintance with it being mediated by dubious perceptual processes, then the existence of our sense-organs must also be in doubt, and therefore the nature of the perceptions from which we started. This is not just difficult but also disorientating, though we always have to act, as Hume would say; but in the meantime we may pause in our struggles, ear flat against the water, and listen for aftershocks as a new world forms from the disintegration of the old.

Finally, there are other minds and the tension between self- and other-ascription of mental states, together with discord between conflicting aspects of our understanding of other people. We track the progress of tears on the cheeks of an individual, who we can see is upset

or in pain, our response to which is to comfort him or her.[2] But always, unless we are beside ourselves, the observer in us remains detached, kept separate by acquaintance only with this other person's appearance and behaviour, her conscious mind wholly inaccessible and her body as alien to us as it would be to her, apart from its surface features. It is essentially these, especially her physiognomy, which exhibit her to us as a person, and this induces in us that crucial sense of her consciousness displayed, of mind personified in the face it inhabits; hence the feeling we have of being in the presence of another human being. Her face is a stage, then, on which her inner life is enacted; but also, again opposing one picture to another, it is as if that stage is a platform, the actor sizing up the audience, her eyes trained on other people like binoculars, and on her surroundings, from which the visual images are transmitted to an occupying mind. Again, then, we may freeze in the headlights of opposing views of this person, stalled by an impulse, immediately checked, to peer deep into her eyes as if through a two-way lens, to the conscious thing behind them. Or, there again, we may yield to that impulse, courting embarrassment if with a stranger and disappointment in any case, for we cannot see into a person's eyes, not even into our own in a mirror, or not without an ophthalmoscope, and if we could there would not be an inner being staring back, but only ocular structures connected to the brain by an optic nerve.

Enough has been said to point to what is unaccountable in the nature of things, though there are other cases in the queue, for no mention has been made of the fugitive nature of the concept of personal identity: such a strong sense of self, but impossible to locate it except as a grammatical subject. And, too, there are deep forces in the world which push against intention concepts, upthrusting them into the philosophical realm, where we may ask how it could be that the age of the earth can be grasped in a minute or two, that of the universe, which is three times older, not taking us even a moment longer to understand. There is, after all, the inescapable fact of the epistemic sterility of each moment, for if I now close my eyes I see what is always there when they are open: that this world of information has shrunk not even to virtually nothing but to nothing at all; for I cannot capture the moment without going beyond it, not even if I say that my mind is blank. And yet, even in this empty moment there is so much that I know; not just that my eyes are closed – though how could I possibly know *that* – but a whole universe of knowledge in a space between dispositional and actual. Belief and understanding must be transtemporal – but how could that be and what does it mean?

---

[2] From this point on I shall use, as the mood takes me, either 'him' or 'her', together with its cognate forms.

One could go on, but better, perhaps, to tease out some of the significance of all this for the problems which I have undertaken to resolve, in particular those of induction and our knowledge of the past. Accordingly, I shall now try to exhibit epistemological scepticism as arising from the clash of concepts reverberating into philosophical consciousness; and, too, I shall ask whether by refuting the sceptic we may expect to resolve the conflicts which generate the tsunami on which he rides. To begin with, if a paradigm of the physically real is taken to be a presently observed solid object, then the corresponding observation report is in its turn taken to be paradigmatic of the kind of description to which a great deal of certainty may attach. Thus, it seems to provide a touchstone against which the sceptic can unfavourably compare more adventurous empirical statements, such as those which convey predictions or assert the occurrence of a past event. What I try to show in the main body of this work, however, is that his arguments are such as to commit him to a rejection of the epistemic validity of observation statements, a position which I defend in a variety of ways. One of them concerns the fact that it takes time to utter any statement, the sense of which unfolds with the utterance, at which point the object supposedly pinned to a present moment has now shifted into the past, knowledge of which, by the sceptic's own lights, is always open to question. Clearly, if this argument is valid then it applies equally to any set of sense-data statements to which it might be claimed that an observation report is reducible. Thus, the sceptic loses his touchstone at all levels and cannot even use the immediately given as a point of certainty from which to challenge our knowledge of the past or the future.

Given these findings, let us now ask again whether, if there really exists a hidden nature of things, and if its effects are manifested in the discordant character of the world as we encounter it, these effects continue to be exerted even when sceptical problems in the theory of knowledge appear to be resolved, thereby imposing constraints on the significance of any solution. In providing an answer, we need only complete the anti-sceptical argument in part just outlined, which in concentrated form runs as follows: the uttering of any statement about a present object, say this pebble, occupies a period of time, during which the pebble has aged slightly. Strictly speaking, then, such existential or descriptive statements refer to a past event. If this is hard to see, it is because of continuity of resemblance, in this case between successive moments in the life of a pebble. Clearly, such statements must be targeted by the sceptic about the past. But now, the reference he thereby makes to them must also be targeted; therefore, scepticism about the past is self-refuting. Arguments of this kind are used in the main body of the work, my point being that they do nothing to moderate the paradoxical character, as it appears to me, of the way in which physical reality is experienced.

Unlike the sceptic, I do not have to withdraw from such experience, which I fully embrace as I sit here in my study – but then, the discordance is all the more jangling when I now try to reconcile the solid reality all around me with the flow of my perceptions. It does not seem to me to be the case, then, that harmony can be achieved by refuting epistemological scepticism about the past or about anything else; nor, indeed, by capitulating to it, or by exploiting or resolving problems about sticks looking bent in water, or by reducing one set of concepts to another, usually taken to be more intimately linked to what is real and knowable. Discordance among concepts appears to be fundamental; hence the suggestion that where they conflict they resonate with something deep and inaccessible.

None of this is to deny that in riding out to confront the sceptic I may embolden myself by this positing of a hidden reality responsible for our experience of the world and for the impossibility of comprehending it. The feeling this evokes is linked to the thesis that our cognitive interactions with the world form a system, and in the following chapters I try to exhibit the interconnections which constitute it and to formulate the conditions of discourse which it frames. Reasoning in this way, I am naturally disinclined to construe certain forms of reality as providing an ontological or epistemological touchstone for others; for instance, I easily resist being seduced by behaviourism about other minds, and I find its inconsistencies unappealing. Also, I take issue with the Wittgensteinians on intentionality, and this again is part of the pattern, the equating of meaning with use being rejected in favour of the irreducibility of what it is to mean and to understand. The followers of Wittgenstein or Kripke would argue, after all, that if I repeat a sentence then nothing in the similarity of form, or of mental accompaniment, including mental images or any feeling, as it were, of the meaningfulness of one's words, constitutes their having meaning; therefore, there is no fact of the matter as to what we mean. But this is to assume that meaning is real only if composed of such conscious items as these; and this assumption, which smacks of empiricism to me, is one for which I am unable to find any grounds.

As part of the pattern, then, I reject what I take to be an unduly restrictive ontology, in this case concerning semantic content, which I take to occupy a category of its own. This is not to imply anything Platonic, but only that we cannot speak of what it is to mean or to understand without employing semantic or cognitive concepts. That, at least, is what I maintain in the body of the text, though I permit myself to add here that if meaning and understanding are irreducible this raises the possibility, precluded by the thesis equating meaning with use, that there may be hidden depths to them. Our concepts are determined by

what we mean and understand, after all, and philosophically interesting conceptual conflicts, or so I have conjectured, may indicate the presence of hidden structures of reality.

That said, speculation of this kind is anathema to some philosophers, and A.J.Ayer, for instance, would suspect metaphysical contamination of the present academic work. It might be thought, however, that to agree with him is to cleave to too strict a standard of intellectual rigour. One does not automatically condemn as metaphysics, after all, the positing of a non-physical reality as the source of one's perceptions, a hidden something known only by the world it presents to us through our senses. Support for this view is then drawn from the wealth of argument on which scepticism about the independence and objective status of the physical world is based. In rejoinder, it is usual to point out, as Ayer (1973a, chapter 4) does, that if all we know, or could know, of this occult region of the real is its observable effects, then we have no reason to postulate its existence. The point to be made is just that the view in question is not self-evidently false or beneath consideration – hence its being suggested that I appropriate it in defence of my own appeal to what is mysterious.

There are two difficulties here, the first of which is that scepticism about the physically real applies to perception itself, to the existence of the senses, a consideration to which the rational response, as remarked upon earlier, is to feel completely out of one's depth, so that one's neat picture of a parallel world slips out of one's grasp into deep water. Secondly, such a world is too much a construct of this limited form of scepticism, whereas my concern is with a wider range of issues and a more inclusive underlying reality. The reason I wish to spread my wings in this way is that the notion of an all-encompassing extramundane reality follows naturally from that of a system characterised by interconnectedness, both of concepts and, concomitantly, of different areas of scepticism. There are, for instance, the connections by which the sceptic about induction is committed to global scepticism.

As already implied, speculation about a reality behind the one with which we are familiar plays little part, or only very indirectly, in my attempted refutation of the sceptic, and neither does it figure in my development of the view that reasoning about the world forms a system characterised in certain ways. It is an offshoot of these but does not itself contribute to them; and this is just as well, perhaps, when one considers that the secret nature of things, if there is one, is wholly unknowable in itself, hard as it is to gain purchase on this fact and to keep faith with it. One temptation, I suppose, is to pin one's hopes on revolutionary insights or discoveries in science, such as to transform our view of the world and offer a glimpse into the unobservable reality which underpins it. This, however, would seem to be a delusion, and

it is no good, either, expecting technologically superior extraterrestrials to be able to rescue us from our unenlightened state. According to Science journal, reported in the Guardian of 28th October 2010, there may be 50 billion Earth-like planets in our Milky Way galaxy, which is to say that terrestrial philosophy departments may not be alone in the universe. But our counterparts on these exoplanets will face the same epistemological problems as we do, and will speculate in the same way, and to no avail, about hidden reality. The galaxy may be teeming with philosophers, not one of whom, really, has any idea what on earth is going on. Here on Earth, meanwhile, I should like to propose a solution to the problem of induction and that of other minds, which at least will serve as a distraction from unanswerable questions concerning an unfathomable mystery.

Chapter 1

The Sceptical Problem and Logical Probability Theory

Outline of the problem

Perhaps I may begin with an outline of the problem of induction from the point of view of David Hume, the philosopher widely regarded as the progenitor of the problem in its modern form. He starts from the empiricist premiss that our concepts of the physical world are derived from experience, or in his own words that impressions give rise to ideas. Thus the notion of a causal relation originates in our experience of the constant conjunction of events, which we associate through relations of resemblance. This is part psychology and part philosophy, for Hume would insist that our understanding of the world *must* have its source in experience, a point of view which is opposed to epistemic rationalism. He would say that what we know of events in the world, or the world beyond our immediate perceptual horizon, derives from our past experience of similar events and that we seek to justify it in terms of them.

Issuing a challenge as to what such a justification could amount to, Hume now argues that inductive inference is non-demonstrative. If by way of example we take my belief that the sun will rise tomorrow, and if in support of it I cite the fact that the sun has risen daily since time immemorial, or at least since I was a boy, then it is clear that in this case evidence does not entail conclusion, for the failure of the sun to rise is at least conceivable, and no contradiction is involved in stating both that the sun has always risen in the past and that it will rise no more.

Is it not possible, however, that the appeal to past regularities could take some other form? Addressing this question, Hume first points out that there is no relation of natural necessity in which factual inference from observational premisses may be grounded. Such a relation does not readily spring to mind when predicting a sunrise on the basis of previous ones, but if we think of mechanical effects, for instance the deflection of colliding snooker

balls, then perhaps the impression we have is one of power and force, which we could then intellectualise as natural causal necessity. Hume's point is that this is an empty notion which contributes nothing to the bare facts of constant conjunction and regularity, there being no significant difference in this respect between indirect causation, as with the regular appearance of the sun, and the direct effect of one object impacting another. It follows that the question about justifying inductive inference still awaits an answer.

One possibility, according to Hume, is that we may appeal to a general inductive principle, that of the uniformity of nature, which states that 'instances of which we have had no experience, must resemble those of which we have had experience, and that the course of nature continues always uniformly the same' (2003, p.61). A question now arises as to the epistemic status of the principle itself, which Hume denies is a logical principle on the grounds that we can at least imagine enormous upheaval in all that we take for granted in the world, which would then become an alien landscape of inexplicable events and cognitive disarray. If the principle cannot be formally demonstrated, then Hume would argue that its validity must depend on contingent facts about the world, which is to say that the evidence in its favour is just that it has proved effective in the past, on which basis we should adhere to it in the future. But this itself is an inductive argument; therefore, it cannot be used without circularity in the justification of inductive inference. Besides, there are those who object to the very idea of a principle of uniformity, which they regard as being deeply problematic, not least because it seems to reduce to the regularities to which it is meant to give epistemic authority. However that may be, the conclusion to which these arguments tend is that we cannot justify even our most entrenched beliefs about unobserved features of the physical world, let alone our more tentative opinions as to what is likely to happen.

That, in essence, is Hume's sceptical thesis about induction, a salient feature of which is that it falls short of outright rejection of all that we claim to know, depending as it does on our knowledge of the past and of presently observable events. There is, as I shall try to show, a key question as to whether restricted scepticism of this kind can be coherent, and a further question as to whether unrestricted scepticism makes any sense at all. We shall seek answers at a later stage, but my present concern is with rejoinders to Hume that do not challenge him in this way, which would seem to be the case with all of them, and in particular I intend to focus on probability solutions to the problem of induction. This class of solutions makes use of applied probability theory; as such, it may be understood in the wider context of connections between necessary and contingent, which we may illustrate first of all in terms of geometric prediction. If I draw a right-angled triangle of sides three and four units, then I may

predict with certainty that the hypotenuse will be of length five units, unless Pythagoras made a dreadful mistake all those years ago. Setting aside this possibility, it is arguable that my prediction has logical force, so that a question immediately arises as to whether I have just provided a counter-example to Hume, thereby bringing this thesis to an abrupt end.

That question also arises with the following application of probability theory. Suppose that I draw at random a single ball from a bag containing ninety-nine black balls and a white ball, such that all alternatives are equiprobable; then very likely the ball I have drawn will turn out to be black. If to reason in this way is to appeal to a probability relation of partial entailment, as some theorists believe, then I have logical support for my expectation as to the colour of the ball; yet according to Hume I have no grounds at all for expecting one colour rather than another.

Keynes' probability theory

Clearly, we cannot dismiss out of hand attempted probability solutions to Hume's problem. To be able to do justice to these attempts, we need to examine them in detail and to familiarise ourselves with the philosophy of probability theory, much of which concerns the logical status of the connection between pure and applied aspects of the probability calculus. To that end, I shall now consider the theory developed by J.M.Keynes, who is one of the originators of what has come to be known as the *a priori* or logical theory of probability. The classical theory was *a priori* in the sense that it applied the principle of indifference to alternatives derived formally rather than by empirical investigation. With Keynes, too, a similar use is made of that principle in a modified form, but in other ways his views are quite distinct and deserve to be treated separately. For him probability is a logical relation between propositions, the symbolic expression for which is a/h, which reads 'the probability of a, given h'. This conditional formulation reflects his belief that 'it is without significance to call a proposition probable unless we specify the knowledge to which we are relating it' (1973, p.4). Despite appearances, Keynes denies that his theory is subjective: 'A proposition is not probable because we think it so. When once the facts are given which determine our knowledge, what is probable or improbable in these circumstances has been fixed objectively, and is independent of our opinion.' (Ibid)

Not only has it been fixed objectively but also, since probability is a logical relation between evidence and conclusion, it follows that the effect of new evidence is to change but not to invalidate the old conclusion. Thus:

As our knowledge or our hypothesis changes, our conclusions have new probabilities, not in themselves, but relatively to these new premises. New logical relations have now become important, namely those between the conclusions which we are investigating and our new assumptions; but the old relations between the conclusions and the former assumptions still exist and are just as real as these new ones. (p.8)

In symbols, if evidence and conclusion are such that $a/h = p$, and if new evidence h_1 is considered, then it may be that $a/hh_1 = p_1$, but it is still true that $a/h = p$. [a/hh_1 is the same as $P(a/h \cap h_1)$, which stands for the probability of a, given h and h_1.]

An example of what he has in mind can be given in terms of tossing a coin. Suppose that my initial estimate, if I have no reason to choose between heads and tails, is that the probability of heads with this particular coin is half. Then Keynes is saying that the logical relation between this evidence and this conclusion continues to be valid even if I toss the coin a hundred times, its landing heads each time, and reach the new conclusion that the probability is very much greater than half. Since probability is a logical relation between premiss and conclusion, it follows that the probability of heads on the original evidence is half and that the new probability of heads on the additional evidence is greater than half. This is to be contrasted with the view that the new evidence indicates, very strongly in this case, that the original probability assessment was incorrect and unjustified.

It might be thought that this contrasting point of view has a lot to recommend it; but my task at the moment is to outline rather than to criticise Keynes's theory. Included in it is the claim that it is not possible to provide a definition of probability as a logical relation, 'unless it contents us to define degrees of the probability relation by reference to degrees of rational belief' (p.8). If probability as a logical relation is indefinable, this raises a question as to how we acquire knowledge of it. The answer, according to Keynes, is that we apprehend it directly through intuition. And what our intuition reveals to us is that probability may be qualitative as well as quantitative. For instance, 'We are out for a walk – what is the probability that we shall reach home alive? Has this always a numerical measure? If a thunderstorm bursts upon us, the probability is less than it was before; but is it changed by some definite numerical amount?' (p.31) This would seem to imply that there is only one sense of probability and that whether or not it has a numerical measure will depend on the nature of the evidence. A numerically precise judgement is possible if conditions are such as to license the application of the principle of indifference. This principle is so important to Keynes that he devotes the whole of a chapter to a discussion of it. His approach is to criticise

its unfettered use and then to suggest restrictions designed to obviate the worst of the criticisms.

What these restrictions amount to is that the principle of indifference should be used in such a way that the alternatives over which it operates are symmetrically related to the evidence and ultimate in the way in which they are selected. For instance, if all that is known about two balls in a bag is that they are black or white, and if we consider the possible colour ratios, then the resulting alternatives are two blacks, two whites and black and white, or *bb*, *ww* and *bw*. If these alternatives are regarded as equiprobable with a probability of one third, and if a conditional probability formula is used, then we arrive at a probability of $\frac{2}{3}$ that the second ball will be black if the first is black and is not replaced.[3]

According to Keynes these alternatives are not symmetric with respect to the evidence, as can be shown by treating the balls as individuals and reviewing the alternative *bw*. This comprises the case in which the first individual is black and the second white and that in which the first is white and the second black. Since these are separate cases, the alternatives are not symmetric with respect to the evidence if there are taken to be three of them. Symmetry is achieved if the four alternatives *bb, ww, bw* and *wb* are considered, and if they are regarded as being equiprobable, in which case they are said to be the colour 'constitutions' of the balls. It then follows that the probability that the second is black, given that the first is black, is $\frac{1}{2}$. The four alternatives are also ultimate, so that they satisfy his second condition.

To highlight this second condition, let us consider the example in which all I know about a ball in a bag is that it is of uniform colour. I could argue that there is nothing to choose between its being red and not-red, therefore they are equiprobable. What is wrong here, says Keynes, is that not-red may be regarded as a disjunction of alternatives of the same kind as red; therefore the contradictory of red is not equiprobable with red, though its contraries are. So it would seem that these two conditions for the application of the principle of indifference remove many of the inconsistencies arising from its unrestricted use.

What is still permitted, however, is the application of the principle to those cases in which I am able to categorise the features of interest without knowing what their proportions are. Just now, for example, we considered the case of two balls in a bag being black or white with nothing known about whether they were two blacks, two whites or a black and a white.

---

[3] $P(b_2/b_1) = \dfrac{P(b_1 \cap b_2)}{P(b_1)} = \dfrac{\frac{1}{3}(1)}{\frac{1}{3}\left(1+\frac{1}{2}\right)} = \dfrac{2}{3}$

Keynes would say that we have no reason to choose between black and white as the colour of the first ball, therefore these colours are equiprobable with a probability of half. This contrasts with an opposing view to the effect that if the proportions are unknown, then we do not have enough evidence on which to base a probability judgement.

Let us now turn to the connection, on Keynes' account, between *a priori* probability and experience. We may start by examining his views on the independence of events with respect to their probability. If you look at a probability question from 'A' level statistics, you will see that when two events are said to be independent this is taken to be an instruction to use a particular formula, which I shall call the product formula, in calculating the probability that both of them will occur. The student is told, for instance, that the toss of a coin is an independent event, and this is taken to mean that if the probability of heads is p, then the probability of n heads is p^n. As regards the value of p, it is usually said that the coin is a fair coin, from which it follows by definition that $p = \frac{1}{2}$.

When 'fair' and 'independent' have the force of instructions their meaning is clear, and for the student the question of defining them does not arise. For Keynes, a definition is thought to be necessary and is given as follows: if $a_1/a_2 h = a_1/h$ and if $a_2/a_1 h = a_2/h$, the probabilities a_1/h and a_2/h are independent. If they are, then it follows immediately that $a_1 a_2/h = a_1/h \times a_2/h$, and this is what I shall refer to as the product rule.

Armed with this definition, Keynes now stresses how important it is to be clear about the nature of independence if mistakes in reasoning are to be avoided. In line with his thesis that probability is a logical relation between propositions, he says that independence is a property of propositions or arguments rather than of events. If we ascribe it to events, we are liable to assume that the independence in question is causal. We would then say, for instance, that the separate tosses of a coin are independent, since the result of one toss does not causally depend on previous results. On this basis we would then apply the product rule, as the student is instructed to do by being told that the toss of a coin is an independent event.

For Keynes, however, what is important is 'whether there is any *known* probable connection, direct or indirect. A knowledge of the results of other tossings of a coin may be hardly less relevant than a knowledge of the bias of the coin; for a knowledge of these results may be a ground for a probable knowledge of the bias'[4] (p.182). If we assume that 'bias' depends on relative frequency, then what Keynes has in mind is that the separate tosses of a

---

[4] In this thesis italicized words in quotations belong to the quoted passage.

coin may be causally independent but they need not, to use his phrase, be independent for knowledge. This distinction is important to him because he conceives of probability as being relative to evidence, which is to say to knowledge, so that knowledge of previous results of tossing a coin may change the probability of heads or of tails on the next toss.

Returning to this point, he says:

An excellent and classic instance of the danger of wrongful assumptions of independence is given by the problem of throwing heads twice in two consecutive tosses of a coin. The plain man generally assumes without hesitation that the chance is $(\frac{1}{2})^2$. For the *a priori* chance of heads at the first toss is $\frac{1}{2}$, and we might naturally suppose that the two events are independent, – since the mere fact of heads having appeared once can have no influence on the next toss. But this is not the case unless we know for certain that the coin is free from bias. If we do not know whether there is bias, or which way the bias lies, then it is reasonable to put the probability somewhat higher than) $(\frac{1}{2})^2$. (p.187)

There are difficulties here, one of which concerns Keynes' use of the notion of equipossibility. The fact is, after all, that if I have nothing to choose between heads and tails on a single toss, so that they are equipossible, then I am in the same initial state of ignorance with regard to the four alternatives, in terms of constitution, arising from a double toss of the coin[5].More generally, tossing the coin *n* times would yield 2^n alternatives, each of them equipossible in the sense that I have nothing to choose between them before tossing the coin. According to the principle of indifference, this should mean that they are equiprobable, in line with what Keynes believes that the plain man would say; but he rejects it on the grounds that heads at the first toss, or the first few tosses, is evidence that the coin is biased, or in other words that heads is likely to be more frequent than tails. It would seem that he wishes us to believe that in order to decide now, in advance of any information, whether there is anything to choose between the alternatives, I have to ask myself what I would expect the second result to be on the basis of the first. Thus, if the first is heads, I should expect the second to be heads rather than tails; therefore – or so we are to believe – two heads are initially more probable than a head followed by a tail, just as two tails are more probable than a tail followed by a head. Clearly, this cannot be correct, for it would mean that despite our ignorance as to the relative frequency of heads and tails, we are somehow privy to the *a priori* knowledge that some alternatives are more likely than others.

Now common sense would suggest that such alternatives are indeed equipossible, though not necessarily equiprobable – or not if one accepts that if nothing is initially known about the relative frequency of heads and tails, then their probability cannot be estimated.

---

[5] Or the three alternatives in terms of ratio.

Keynes equates the equipossible with the equiprobable, which is why he denies that these alternatives are equipossible, for this would entail their equiprobability. They cannot be equiprobable, he thinks, given that relative frequency is evidence of bias, for there would then be a problem with applying the probability calculus. To show this, he lets h_1 = 'heads at first toss', and h_2 = 'heads at second toss'. Then ' $h_1/h = h_2/h = \frac{1}{2}$, and $h_1 h_2/h = h_2/h_1 h \times h_1/h$. Hence $h_1 h_2/h = \{h_1/h\}^2$ only if $h_2/h_1 h = h_2/h$, *i.e.* if the knowledge that heads has fallen at the first toss does not affect in the least the probability of its falling at the second' (p.188). But it does affect it, he says, and that is why the initial probability of heads twice is greater than half squared. These equations are based on Keynes' own system of mathematical probability, but they have their counterpart in Bayes' Theorem, which in its simplest form, and in modern notation, states that $P(a/b) = \dfrac{P(a \cap b)}{P(b)}$. This may be expressed as $P(a \cap b) = P(a/b) \times P(b)$, which corresponds to $h_1 h_2/h = h_2/h_1 h \times h_1/h$. Bayes' Theorem has a standard use in which the known prior probabilities $P(a \cap b)$ and $P(b)$ yield the new probability $P(a/b)$, as would be the case with the coin example if the four alternatives were taken to be equiprobable, for we should then have the following:

$P(H_2/H_1) = \dfrac{P(H_1 \cap H_2)}{P(H_1)} = \dfrac{\frac{1}{4}}{\frac{1}{2}} = \dfrac{1}{2}$. Here H_1 and H_2 are the events that the coin lands heads at the first toss and heads at the second toss respectively. In general, if the coin is to be tossed n times, always landing heads, and if the 2^n alternatives are taken to be equiprobable, then the conditional probability of the $n+1$th toss turning up heads is always equal to half. This, however, makes no sense at all in the present case, for the coin examples are such that nothing is known in advance about the relative frequency of heads and tails. Far from being discounted, an initial sequence exclusively of heads would strongly suggest the hypothesis that the probability of heads was greater than half.

The upshot is not that the product rule cannot be used in coin tossing experiments but that the probability of heads and tails must already have been established in previous trials or taken for granted, as it would be if the assumption was that a particular coin was a fair coin. Keynes' difficulty, which seems to me insuperable, is that of having to reconcile the principle of indifference with criteria of probability based on relative frequency. If I am a statistician and I wish to check the hypothesis that a particular coin is a fair coin, let us say, then I subject it to a series of trials in which the coin is tossed and the outcomes recorded, a battery of

statistical tests then being performed on them. None of this need involve the principle of indifference as a logical principle, and I do not have to posit in advance the equiprobability of heads and tails or of the 2^n alternatives in n tosses of the coin.

Keynes, of course, would view frequency considerations in a very different light, as is evident from another example he discusses, this time involving the distribution of black hair and colour blindness in a population. In this example the proportion of black-haired men is $\frac{p_1}{q}$ and of colour-blind men is $\frac{p_2}{q}$, where black hair and colour blindness are assumed to be independent. This is to say that $P(b/c) = P(b)$, where $b =$ black-haired and $c =$ colour-blind. The product rule now follows mathematically, so that $P(b \cap c) = \frac{p_1 p_2}{q^2}$ if the probabilities are identified with the population frequencies. But, says Keynes, the product rule holds '*only* if this is the *actual* proportion in fact of persons who are both colour-blind and black-haired, and that this is the actual proportion cannot possibly be inferred from the *independence for knowledge* of the characters in question' (p.115). In order to evaluate his argument we need a numerical example. Suppose that of two men all that is known is that one is colour-blind and one has black hair, possibly the same one. Letting m_1 and m_2 be the two men, we have the set of alternatives:

$m_1 b \ m_2 c$
$m_1 c \ m_2 b$
$m_1 bc \ m_2$
$m_1 \ \ m_2 bc$

Their meaning will be made clear if I translate the last line: m_1 by itself means that the man m_1 is neither black-haired nor colour-blind; $m_2 bc$ means that the man m_2 has both features. If the alternatives are equipossible, which they are, and if they are therefore equiprobable, according to Keynes, then with regard to each of the men $P(b) = \frac{1}{2}$, $P(c) = \frac{1}{2}$ and $P(b \cap c) = \frac{1}{4}$. Mathematically, then, the product rule holds in this case, since

$P(b \cap c) = P(b) \times P(c) = \frac{1}{4}$.

Now Keynes' argument is that the product rule *must* hold, since the two features are independent in the required sense. Therefore, if it can be shown that the rule does not always hold when probability is based on frequency, then this is an argument against their being connected in that way. In the present instance he would have to say that the product rule

holds only if one in four of the men are both black-haired and colour-blind, a possibility which may safely be ruled out on the grounds that there are only two men.

A weakness in this reasoning is that it can easily be stood on its head, for if the principle of indifference is incompatible with frequency considerations, then it is arguable that the fault must lie with the principle, given how counter-intuitive it would be to divorce probability from frequency. If, in the present case, we assume that the marriage is intact, then $P(b) = \frac{1}{2}$ and $P(c) = \frac{1}{2}$. But it makes no sense to say that $P(b \cap c) = \frac{1}{4}$, for this would mean that that the probability of a man selected at random being both black haired and colour blind is a quarter, so that it is rational not to expect this combination of features, whereas in fact we have no idea whether to expect it or not.

The point being made is just that equipossible does not entail equiprobable, as will perhaps be made clearer by the following example, this time involving balls in a bag where the proportion of colours is unknown. Suppose that a bag contains three balls each of which is red, white or blue, and that one ball is to be drawn at random. Then it follows from the principle of indifference that the three equipossible colours are equiprobable, against which it could be argued that this is the case only if there is one ball of each colour. Or, it could be argued that if we now take one of the colours, say red, then by the principle of indifference the probability of not-red is two-thirds, so that I am rationally committed to expecting a ball drawn from the bag to be blue or white. But this implies a belief on my part that the bag contains at most one red ball, and this is a belief for which I have no justification, given that I have no idea which colours the balls are, and that for all I know they could all be red. This would seem to establish the point.

In insisting with regard to the above example that the proportion of colours has to be known before a probability judgement becomes possible, I do not mean to endorse the frequency theory of probability, one objection to which is that grasp of probability implies differential rational choice among unequal probability values. If I know that a bag contains nine black balls and a white ball, each equally likely to be drawn, despite which I expect to draw a white ball, then it has to be that my grasp of the notion of probability is deficient in crucial respects. There is more to it than frequency, then, and criticisms of the principle of indifference also have a wider scope, as the following discussion will show.

It concerns Keynes' position on the question of inverse probability in cases involving unknown proportions. With regard to balls in a bag being black or white, the proportions unknown, he says that if equiprobability is ascribed to constitutions as opposed to ratios, this

being what he favours, then, 'the drawing of one ball and the resulting knowledge of its colour leaves unaltered the probabilities of the various possible constitutions of the rest of the bag' (p.53). Take, for instance, the case in which each of two balls in a bag is either black or white, one of them being drawn without replacement and found to be black. Let b = black, w = white, so that the alternatives are bb, bw, wb and ww; let b_1 and b_2 = first ball black and second ball black. Then the conditional probability formula, in modern notation, gives:

$$P(b_2/b_1) = \frac{P(b_1 \cap b_2)}{P(b_1)} = \left(\frac{\frac{1}{4}}{\frac{2}{4}}\right) = \frac{1}{2}.$$ In the same way, it will be found that with n balls

$$P(b_{n+1}/b_1 \cap b_2 \cap ... \cap b_n) = \frac{1}{2}.$$

This is reminiscent of the implausibility of what Keynes has to say about tossing a coin of unknown bias. In both cases if constitutions are deemed to be equiprobable, the results are counter-intuitive, in the present case because if most of the balls drawn from a bag are black, then this will be the colour expected of the next ball – unless, that is, one believes after reading Keynes that the probability is always half.

Having made that point, we may now turn to what seem to be significant differences between the two cases. A coin-tossing experiment, statistically interpreted, may reveal not just whether heads or tails is more to be expected at the next toss but also a specific probability value, or range of values, given by the frequency ratio of heads and tails, where this may be regarded as exhibiting a propensity or disposition of the particular coin. On the other hand, the result of drawing without replacement balls from a bag, though it may be indicative of the colour to expect at the next draw, is not so easily associated with numerical probability and relative frequency. If constitutions are allowed to be equiprobable, then conditional probability values are static, as in the above example, whereas a lengthening run of predominantly black balls would be regarded as increasingly strong evidence for the next ball also being black.

If, despite Keynes' criticisms, we allow equipossibility to attach not to constitutions but to ratios, then the good news is that in the present case probability values increase according to strength of evidence. If we suppose that a ball is drawn and we ask about the next ball, letting the ratios be given by bb, bw and ww, then

$$P(b_2/b_1) = \frac{P(b_1 \cap b_2)}{P(b_1)} = \frac{\frac{1}{3}(1)}{\frac{1}{3}(1 + \frac{1}{2})} = \frac{2}{3}.$$ More generally, if n balls are drawn, all turning out to be black, then the probability of the $n+1$th ball being black, given this application of the

principle of indifference, is $\dfrac{n+1}{n+2}$. This is a form of Laplace's Rule of Succession, but the bad news is that it is totally unsuitable as a probability metric, as is easily shown by consideration of the present example. We are asked to believe that whichever colour the first ball is, the second has a two-thirds chance of being the same; but the truth is that we have no idea which of the three alternatives will be the case.

If we now ask how such probability ascriptions are to be verified, then one answer is that it would require a series of trials in which numerous bags with two black or white balls were sampled, one ball being drawn and then the other. Clearly, however, this would be a futile experiment, for it would be purely adventitious whether or not two thirds of the bags contained balls of the same colour.

Keynes, of course, does not accept the succession rule, for he rejects ratio in favour of constitution. However, a significant point can be made if we note that with both ratio and constitution the outcomes are the same if one calculates the probability of the first ball having a particular colour, which according to Keynes would be 50% if there were two possible colours. If there are any number of balls, and if n colours are given as possible, then the corresponding probability is $\dfrac{1}{n}$ for both ratio and constitution. Consider, then, the simplest case, that in which a bag has a single ball, and this time let the alternatives be numbers, not colours, so that all that is known is that the ball is marked with a number, say an integer in the range 1-100. Then Keynes would insist that since there is nothing to choose between the numbers it would be rational to expect not-1 as opposed to 1, and furthermore that $P(not\text{-}1) = \dfrac{99}{100}$, just as $P(1) = P(2) = ... = P(100) = \dfrac{1}{100}$.

Now, it is arguable that for these probability values to be correct, or even for them to have empirical meaning, they need to be construed in terms of frequency, except that we now have a question as to what the empirical application of $\dfrac{1}{100}$ as a frequency ratio would be. One possibility is that the frequency is that of each number in a series of trials in which numerous bags have their single ball drawn from them. But this is even more absurd than in the last example, for it would require of numbered balls in bags that each of the numbers between 1 and 100 was more or less equally distributed.

Why is it, then, that it seems counter-intuitive to deny that the probability of not-1 is ninety-nine in a hundred, or at least that it is much greater than that of 1? The reason, I

suspect, is that we confuse the probability of an event with that of guessing correctly with regard to it. This is shown in the present case by imagining circumstances in which it would be doubted whether these probabilities obtained. Suppose, for instance, that a dodgy character sidles up to us in a pub, explains that the bag she is carrying contains a ball marked with a number between 1 and 100, invites us to bet that the number is not 1, and offers us what would be very favourable odds if we agreed that $P(not\text{-}1) = \dfrac{99}{100}$. Clearly, we would reject this offer, otherwise there could be no explaining how we had survived long enough to be of drinking age; and it is obvious, too, that our distrust would extend to any situation in which the number or its complement was chosen for us. But if we take the one hundred alternatives to be equiprobable only if we are free to choose the numbers, then this is a standard case of frequency-linked probability, in this instance the probability of guessing correctly, which would be one in a hundred for whichever number was randomly selected, and ninety-nine in a hundred for its complement.

Are there logical probability relations?

Keynes' position, as we have seen, is that probability is evidentially linked to frequency but in itself is a logical relation between propositions, one that in numerical cases gives logical force to the use of the calculus of chances to derive probability values from the theory of arrangements. He allows for such relations in the absence of these values, as illustrated, one might think, by the relation of inductive inference between drawing only black balls from a bag and predicting that the next ball will be black. Here, if anywhere, one might expect to be able to intuit the logical probability relation, given that the machinery of the probability calculus is at rest. But all that one's intuition delivers, or so it seems, is a firming up of belief as favourable instances accumulate, this being expressed as the proposition that it is increasingly likely that all the balls in the bag are black, or that at least one of those remaining will be. The point is that nothing in this process of reasoning suggests the intuiting of a logical probability relation.

There is, if a later discussion is on the right lines, strong support for such reasoning in the form of the probability of guessing correctly, itself connected with the frequency component of what has come to be known as the no-miracles argument. My present concern, however, is with logical probability, characteristic of which, since probability is relative to evidence, is that if $P(a/b) = p$, and if new evidence is brought into the equation, such that

$P(a/b_1) = p_1$, then it remains true that $P(a/b) = p$, which therefore is indefeasible. There are parallels here with geometrical theorems construed as analytically true, thereby acquiring immunity to being falsified by countervailing facts about physical lines and angles in a particular case. What is also acquired is a puzzle as to how it is that the theorems can determine facts about diagrammatical representations of geometrical figures. Since the link between pure and applied probability theory is puzzling in a similar way, it may be that a comparison between the two will throw light on each of them.

Let us begin with the purely mathematical aspects of probability theory, this time in connection with Bayes' theorem in the form $P(h_3/h_1h_2) = \dfrac{P(h_1h_2h_3)}{P(h_1h_2)}$, in which the intersection sign, $\cap$, has been left out, and h_n stands for 'heads at the nth toss' with regard to a series of trials with a fair coin Here, too, there are deductive relations among the three terms, these deriving from the rules by which the probability formula is defined. To see this, we need to set out the alternatives, letting t stand for 'tails':

$h_1h_2h_3$
$h_1h_2t_3$
$h_1t_2h_3$
$h_1t_2t_3$
$t_1t_2t_3$
$t_1t_2h_3$
$t_1h_2t_3$
$t_1h_2h_3$

By definition $P(h_3/h_1h_2)$ is the ratio of the single alternative $h_1h_2h_3$ to the total of alternatives containing $h_1 \; h_2$. From the list, we see that $P(h_3/h_1h_2) = \frac{1}{2}$. Also, $P(h_1h_2h_3) = \frac{1}{8}$ and $P(h_1h_2) = \frac{1}{4}$. So $P(h_3/h_1h_2) = \dfrac{P(h_1h_2h_3)}{P(h_1h_2)}$.

This, on one interpretation, is mathematical probability operating over arrangements, and the formula may then be viewed as analytically true. The reason is that the right hand side unpacks what is contained by definition in the left hand side. To establish a connection with the real world we need to give empirical meaning to the alternatives, which then stand for head and tail sequences, and we need to regard them as being equiprobable. The problem for Keynes, as we have seen, is that they are equipossible, therefore equiprobable, only if the coin is already taken to be fair, since the formula always gives the conditional probability of

heads as half, however long the previous run of heads has been. But all that can be meant by 'fair', or so it seems, is that the alternatives are equipossible and therefore equiprobable. For Keynes it cannot be a reference to frequency, not if probability is logical and *a priori*, and if he tries to take account of frequency as evidence of bias, then the coin is no longer regarded as a fair coin; hence the difficulties that we have seen that he encounters.

That logical probability says nothing about the world is a proposition embraced by many commentators on *a priori* probability theory, not all of whom condemn it on that account. A.J.Ayer does condemn it, and in *Probability and Evidence* he suggests that very often there is a failure to distinguish between pure and applied probability. Since the calculus of chances is purely mathematical in character, he says, the probability values derived from it do not determine what it is rational to believe about actual events, but refer only to relations of ratio and proportion among abstract possibilities. It is therefore quite separate from facts about relative frequency with regard to dice, coins and games of chance in general, about which he says:

If we are going to apply the calculus of chances to actual games, we have to make the assumption that all the logically equal possibilities are equal in fact, and this of course is not a mathematical truth. It is an assumption to which one has to give an empirical meaning.
    Unfortunately, this is often overlooked. It is a common mistake for those who are writing on this subject to start with a well-defined notion of *a priori* probability, namely that of the incidence of certain features in the total of all possible mathematical combinations of a given sort, and then go on to talk of its being or not being equally probable that this or that possibility will be realised in nature, thereby using the word 'probable' in a way that their definition no longer covers. (p.30)

Now, a possible rejoinder here is that the use of the calculus of chances very often yields probability results which accord with standards of rational belief and with frequency desiderata, despite all the exceptions to be found if the calculus is regarded as operating according to the principle of indifference, or if that principle is allowed to determine the range of operation of the calculus. According to Roy Weatherford, for example, 'The strongest single criticism against AP [*a priori*] theories is that the number of logically possible outcomes in no way determines the probability of one of these outcomes in the real world – no merely logical division can tell us what will probably happen in the future.' (1982, p.128) He goes on to say that if this objection 'is valid in its general form, that no merely logical division can determine probabilities, it discredits *every possible* AP theory' (Ibid). He points out that the usual response to this criticism – a response he appears to agree with – is that *a priori* probability

Doesn't prescribe to the world because it doesn't say anything about the world. *But* of the many possible a priori probability systems, some will resemble rather closely the relationship

between natural events – just as some systems of deductive logic and some parts of mathematics resemble the structure of the world. When this happens, we have a tool to use in experience – a partial model of the world. If we then wish to know the consequences of a real event, we ask the system for the modelled consequence in deductive logic, mathematics and probability theory. We then correlate the modelled consequences with another real event, and act on that expectation. Evolutionary and pragmatic adjustments in these systems have brought two of them into very close agreement with the world – it is now our task to systematise and refine our probability intuitions to bring them also into agreement with reality. (Ibid)

A difficulty here is that with regard to the theory of logical probability it seems to be suggested both that probability is a logical relation that says nothing about the world and that probability systems may be used as mathematical models from which the chances of actual events may be derived. As a response to criticisms of the theory, this, however, would seem to concede the point of them, for if the systems are useful only insofar as they bring about agreement with reality, presumably by way of relative frequency, then this is to say that the inferences they generate are empirical in character; and none of this would seem to involve a relation of partial entailment between propositions. There are issues here the resolution of which would perhaps be facilitated by a more detailed examination of the link between pure and applied probability.

Suppose, then, that our interest lies in the probabilities of heads and tails for a particular coin, about which we know nothing in advance, so that heads and tails are equipossible. If we say, with Keynes, that they are equiprobable, then this is misleading, for it implies that the coin is a fair coin, and indeed it was argued earlier that the principle of indifference avoids difficulties only if a coin is taken to be fair, which this one is not. That being the case, suppose that we subject the coin to a series of trials in which it is tossed and the outcomes recorded, and that the sequence of heads and tails is such as to support a hypothesis as to their probabilities. Then an important point, already touched upon in the case of what it is for a coin to be fair, is that probability-value ascriptions impose constraints on future relative frequency as truth conditions, a concomitant of which is that those values determine what it is rational to expect when the coin is tossed. One argument in support of this frequency condition may be given in terms of the grasp of probability as a concept. If in the present case I arrive at a probability value of, say, $P(h) = \dfrac{1}{3}$, but at the same time deny that I would expect the frequency of heads to be about a third, claiming instead that it would be half, or nine-tenths, or just that I would expect heads at the next toss, then clearly I do not understand what numerical probability means. If this is correct, then to say that $P(h) = \dfrac{1}{3}$ is to

imply that the sequence of heads and tails will continue to exhibit certain frequency distribution properties if the hypothesis is to be confirmed. But now, the question of what the future will hold if the trials continue is straightforwardly empirical and relies on inference from those already conducted.

Very well, but if we take it that $P(h) = \frac{1}{3}$, then sophisticated mathematics may be brought into play. In general, if a coin is such that $P(h) = p$ and $P(t) = q$, where $q = 1 - p$, and if X is the random variable 'the number of heads in n independent trials', then the probability density function of X is given by $P(X = x) = {}^nC_x q^{n-x} p^x$ for $x = 0, 1, 2, ..., n$. In this general case X follows a binomial distribution with parameters n and p; and in the present case, in which $p = \frac{1}{3}$, we can let $n = 10$ and $x = 5$, say, so that the probability of five heads in ten tosses is given by

$${}^{10}C_5 \left(\frac{2}{3}\right)^5 \left(\frac{1}{3}\right)^5 = 0.14.$$

Now, it is unassailable that in the given conditions and for the given values the required mathematical probability is 0.14, from which it would be rational to expect other than five heads in ten tosses. What this means, however, is just that the epistemic burden is carried not by the mathematics, which is truistic and bears no weight, but by the question of whether the probability values on which the formula operates are correct, the condition for which is that the sequence of heads and tails should continue to satisfy certain criteria. Since one such value is given by $P(h) = \frac{1}{3}$, its truth condition is that there will be a corresponding relative frequency if the trials continue, failing which, other things being equal, it may be concluded that the value is incorrect. At that point, further into the coin-tossing experiment, there will be new frequency data to be added to the old so that a new probability estimate may be worked out. None of this is to suggest that quasi-logical relations are at work, nor does it help us to intuit what could be meant by them. All that can be said, surely, is that Keynes is correct in his claim that probability is relative to evidence, except that this would seem to be common ground between differing probability theories.

If these arguments are correct, then Hume's inductive scepticism has nothing to fear from the claim that partial entailment gives logical force to the principle of indifference, since the relation between equiprobable and equipossible is not a logical one. It is true that the first implies the second, but not that the converse is the case. It seems to me, indeed, that the

principle should be rejected, mainly because, as we have seen, it conflicts with the role that we assign to frequency when applying the probability calculus. Its place would then be taken by probability theory viewed as a branch of applied mathematics, with all that this implies about the network of relations between pure theory and empirical application.

Is geometry synthetic *a priori*?

Applied mathematics has its own problems of analysis, and this reminds me of the promise I made, that I would seek enlightenment in an examination of geometry as lines on paper, using the theorem of Pythagoras as a focal point. Whether the link between pure and applied probability will be thereby illuminated remains to be seen, but the question of *a priori* knowledge of the properties of geometrical figures may in any case be treated separately, given the possibility of discovering counter-examples to Hume.

With Pythagoras, then, the theorem states that in a right-angled triangle the square on the hypotenuse is equal to the sum of the squares on the other two sides. Thus, if the sides are three units and four units, where $3^2 + 4^2 = 5^2$, then the hypotenuse is five units. This theorem has a physical counterpart in the form of a right-angled triangle drawn on paper. It could be argued that if the sides are known to be three and four inches long, then it is deducible that the hypotenuse is five inches long; therefore, a statement to that effect is a synthetic *a priori* statement. An empiricist then rejoins that a mathematical theorem is being confused with facts about lines on paper. To show that there is a serious issue here, all I need do is to present both sides of the debate before attempting to resolve it. Perhaps the most economical presentation would take the form of a dialogue along some such lines as the following, where persons F and D, father and young daughter, are hunched over a school exercise book showing a right-angled triangle:

D: You don't mind helping me with my homework, do you, Dad?

F: No, love, not at all.

D: Would you agree that these two sides of this triangle are three and four inches long and are perpendicular?

F (using a ruler and protractor): Yes, I agree.

D: Then if we apply Pythagoras to get $3^2 + 4^2 = 5^2$, it follows that the hypotenuse is five inches long, this being a deductive conclusion from true premisses.

F: In geometry it's easy to confuse mathematics with physical facts. If a conclusion follows deductively, then it and the premisses are mathematical. The hypotenuse being five inches long is a contingent fact about lines on paper.

D: But Dad, if we believe that this is a right-angled triangle of sides three and four inches, then we *must* believe that the hypotenuse is five inches long.

F: You seem to take the geometrical application of the Pythagoras theorem to be synthetic *a priori*, as if the theorem's logical force somehow leaps the gap between pure and applied; but I would say, following Ayer, that the theorem is analytic, and logically separate from empirical fact.

D: Well, Dad, serious doubt has been cast on the notion of analytic truth, notably by Quine, who argues, as Descartes would put it, that we have no clear and distinct idea of the synonymy on which the notion of the analytic depends, partly because of the way in which the synonymous and the analytic feed into each other. Besides, what makes you think that the theorem is analytic? Can you show that it is?

F: No, and you are right to challenge me, though perhaps I would have more success in simpler cases, and in a way that, *pace* Quine, would be unproblematic. Don't forget, too, that he would question the notion of synthetic *a priori* truth on which you seem to rely. What I can still say, in any case, is that it is an empirical question whether this hypotenuse is five inches long, irrespective of whether the theorem is analytic or whether it is logical, assuming any difference; and also, again following Ayer, that nothing is allowed to count against the truth of the theorem. If the hypotenuse turned out to be other than five inches, I would seek an explanation, but not at the expense of the theorem, which would always be inviolable, the problem lying with its physical application and not with its mathematical validity. I would check the measurements again, confident that they would turn out to be incorrect.

D: But it is precisely because our belief about the length of the hypotenuse relative to the sides has logical force that we check our measurements when there seems to be a discrepancy. Why else would the length of the hypotenuse give us any reason to doubt whether we have correctly measured lines and angles?

F: Well, I suppose our reasons must be inductive. If all the right-angled triangles we have ever encountered have been Pythagorean, then there is no reason to think that this triangle is an exception; but it isn't logically impossible. Keep in mind, after all, that this hypotenuse *looks* about five inches, even before being measured; but just imagine that the sides were much longer, say thirty and forty miles – would you then insist that the hypotenuse must be fifty miles?

D (failing to sense a trap and trying to imagine a ruler that long): Yes. It would be difficult to measure, but it would have to be 50 miles, given sides of 30 and 40 miles and an included right angle.

F (springing the trap): But then you would be mistaken. Because of the curvature of the Earth the triangle would occupy curved, not Euclidian, space, and the hypotenuse would be longer than 50 miles.

D (recovering quickly): But that just means that a different formula would apply, and would dictate, in the same way as the Pythagoras theorem in Euclidean space, the length of one side of the triangle if the other two sides and the included angle were given.

F: What a clever little girl you are.

D: Dad, are you patronising me?

F: No, love, I wouldn't dare.[6]

And so on. Looking back over this dialogue, we can see that it plays over difficult issues, partly because with geometrical figures the abstract and the physical merge together into lines on paper. It is like, and also unlike, the dual-image effect of looking at certain shapes and seeing first one pattern and then another. As one image comes into focus, so it is that the words 'given a side of a triangle three inches long' seem to refer to a physical line on paper. As this image morphs into another, so it is that the physical line appears as no more than a linear representation of an abstract entity. There are parallels here with the manner in which mathematical and applied probability merge together. If a written probability problem uses the language of everyday objects, we read it accordingly; but also we read it as mathematics and our first step is to find symbols on which to operate. But these include letters which themselves are dual-image, since they also represent the original objects.

If we now try to keep these images apart, then one way is to focus on the act of drawing and measuring the sides and included angle, so that we see it as a physical process. If we take the end points of the sides, the connecting hypotenuse will turn out to measure five inches, failing which it is to be expected that a mistake has been made in drawing the triangle. Since all this is attended with a great deal of inductive and perceptual certainty, we need to distinguish it from the logical force of the Pythagoras theorem. Logical force is, in a sense, absolute, but this should not detract from the strength of our beliefs about matters of empirical fact. This is such that the whole of our waking experience is conveyed from one moment to the next by certainties of detail, which familiarity both channels and obscures.

---

[6] The references in this dialogue are to the views of W.V.O.Quine and A.J.Ayer, in particular Quine's *From a Logical Point of View* and Ayer's *Language, Truth, and Logic*.

Consider, for instance, my reaching for this cup without looking, whereupon I bring it to my lips and sip the tea, every step of which is automatic. There is always, of course, the possibility of surprise, as when I lift the cup only to find that it is lighter than it should be. But suppose that I glance inside it, expecting it to be empty, only to find that it is full of tea, which, bewildered and not being house-proud, I try to expel from the cup by turning it upside down, except that nothing comes out. Now imagine the shock of it and the associated cognitive disorientation, an exercise which may suggest that if we focus on logical force, then we underestimate the power of everyday beliefs. It is true that certainty about the length of a hypotenuse is not part of the fabric of everyday thinking; but it occupies a place in the minds of mathematicians and many philosophers, who have every reason to believe that the theorem of Pythagoras is satisfied by all carefully drawn right-angled triangles, just as unsupported tea always falls.

This is not enough to prove that facts about physical triangles are contingent, but we may reinforce the argument by noting that if we measure such a triangle more precisely than with a ruler and protractor, then invariably we find that even if those blunt instruments show it to be right-angled and Pythagorean, it satisfies the theorem of Pythagoras only approximately. This is good enough for illustrative use and also for checking and discovering properties of the triangle which correspond to those deducible from the theorem, despite the fact that the match between them is only approximate. Thus, if the little girl is asked to measure the length of a hypotenuse of what she is told is a right-angled triangle of sides 5 and 12 inches, then a ruler is accurate enough to yield an answer of 13 inches if she assumes that the length will be a whole number of them. What this indicates, given such facts about approximate correspondence, is that the deductive truths of geometry do not yield *a priori* knowledge of the measurement properties of triangles drawn on paper or embodied in some other way. What we should infer from this, surely, is that physical triangles *represent* geometrical theorems. This proposition gains further support if we now consider Ayer's claim that no empirical fact, however anomalous, could be taken to disconfirm the mathematical or logical truths to which it fails to correspond. Thus, if the hypotenuse in the present case turns out to be 14 inches, not 13, then at no point in the process of checking, re-checking and scratching her head will the little girl entertain the possibility that the theorem of Pythagoras is flawed. Besides, in modern mathematics the theorems of geometry have an algebraic counterpart which is wholly abstract in its use of symbols, none of which designates the vertices of triangles or the angles between their sides.

Let us now take it all a step further, this time discussing Ayer's claim in terms of the simpler case of linear inequalities. One example would be the following: if $a > b$ and $b > c$, then $a > c$. Since the arrow symbolises the transitive relation 'greater than', Ayer would argue that the conclusion is contained in the premiss and is therefore analytically true. If the inference is given factual content, as when a, b and c are planks of wood compared with regard to their lengths, then we may ask whether it is necessarily true that a > c if a > b and b > c. According to Ayer[7] it is, but only in the empty sense that nothing is allowed to count as a contrary instance. This becomes clear if we separate abstract inequalities from comparison of unequal lengths, the need for which is apparent from the fact that if '>' means 'greater than', then it may be used in both cases, just as numerals designate both abstract number and physical measure. In the interests of clarity, then, we should say that if the proposition 'If $a > b$ and $b > c$, then $a > c$' is taken to express a logical truth, then its terms are abstract and cannot enter into factual statements. If, on the other hand, we measure planks d, e, f, such that d is longer than e when measured, and e longer than f, which nevertheless turns out to be longer than d, then although this will be entirely unexpected, it has nothing to do with the logical fact that if $a > b$ and $b > c$, then $a > c$.

First steps towards a solution

To leave it at that, however, is to strictly demarcate the logical from the factual, neither of which, thus distinguished, could provide a solution to the problem of induction. As a first step towards that solution, I shall now argue that from the logical possibility of plank f being longer than plank d in a particular case, it follows not at all that a world is possible in which such counter-intuitive results are the norm. To see this, we note first that if a tape measure is used on the planks, its function is to transfer lengths, in this case as a substitute for placing two planks together and aligning their ends, so that a direct comparison can be made. If this is done, there is a point to the exercise only if, as is clearly the case, the planks are individually identifiable and are taken to be stable in size. This is to be contrasted with having a mental image of two lines of unequal length, which then disappear and are replaced by two other unequal lines, such that none of the lines is representative of a physical line with its own identity, any more than a mental image of blue need represent a blue object. That being the case, it makes no sense to speak of a line from one pair being the same as one from the other, or being different in a sense that implies the possibility of being the same. It follows that

---

[7] Again see his *Language, Truth and Logic.*

nothing could count as my having a mental image of two unequal lines, then another mental image of the shorter of the two juxtaposed with an even shorter third line, which in yet another mental image is compared against the longer of the two. It is possible, of course, to have a mental image of three unequal lines, as it is of two, but then it is given in the image itself that one line is longer than the other two, so that no inference is made from two inequalities to a third.

In the actual world of trees and timber, planks in general are taken to have numerical and qualitative identity through time, as also to have constancy of length and other spatial properties, the same being true of the physical environment in which they are measured and compared, this being the essential framework in which the concept of size and change of size has application. Thus, if plank f turns out to be longer than plank d in any particular case, the explanation being that one or more of the planks must have changed its length, then this can be said only if the lengths are now taken to be fixed, at least while f and d are being compared. If the planks are measured, then this has a point only if the tape measure is taken to have uniform gradations and to be stable in that respect. The conclusion to be drawn is that inequality inferences have physical application only if the objects compared are taken to exhibit constancy of size. This is also a necessary condition for the corresponding equality inferences.

What I shall now try to show is that the converse is also the case, so that the necessary relations between such inferences and the world are bi-conditional, which is to say that not only does the world have to be stable if the inferences are to hold, but also that they have to hold if the world is to be stable. For this to be shown, we need to be able to refer to separate observations of the physical world, the separations occurring whenever one's perceptual experience is interrupted, however momentarily, as when we blink while looking at a plank, or look away and then back. Suppose, then, that we look at contiguous planks d and e, seeing directly that $d > e$, then look again and find that e $> d$. Now try to imagine such changes occurring over the whole of our visual field, the process repeated every time a new observation is made. Clearly, this is difficult to imagine, for it would mean, for instance, that every time I looked away from this table, or simply moved my eyes over it, its inequalities of length, breadth and thickness would change, as would its appearance in relation to those parts of the room that frame my field of view, where these, too, would change in relation to the table and to one another. But then, the way in which the world presented itself to me would be so transformed that it would no longer mean anything to speak of a table or a room, or indeed of a physical world at all if these kaleidoscopic changes attended my perception of it.

Now that we have established that a world of incomprehensible visual flux would result from physical inequalities no longer continuing to hold, it should be easy to show that the effect would be the same if inequality inferences were no longer truth-conducive. To that end, suppose that such inferences generally failed, so that if one object was observed to be greater in size than a second, itself observed to be greater than a third, then this last object would not turn out to be less than the first. Then it would have to be that the relative size of the objects had changed, and if this were a universal phenomenon the outcome once again would be a disjointed world of chaotic visual appearances in which none of our physical object concepts could have application. This concludes our proof that the world can be stable if and only if inequality inferences are truth-conducive.

To conclude this chapter, let me now focus on the insights to be gained from our excursion into transcendental arguments. To begin with, I have tried to show that the logical and the empirical do not form a strict dichotomy, for although it is true that nothing counts against the logical force of the abstract inequality inference, such that if $a > b$ and $b > c$, then $a > c$, it is also true that a physical world of persisting objects is possible only if that inference continues to be valid when given concrete application. Concomitantly, this thesis about inequality inference and necessary conditions extends to other transitive relations, including those of spatial position, as when objects or their parts are to the left of one another, or to the right, or above or below, or inside or outside. And, of course, these and other relations may hold between more than three terms, with all that this implies about the increased scope for inference. That is the thesis, and if it is correct then it calls upon *a priori* knowledge of necessary conditions for the existence of a stable world. It may be worth suggesting, too, that a fundamental condition for the existence of that world, or for the possibility of knowledge of it, is that mathematics should apply to it, not only at the basic level at which objects in the world are countable but also at the higher level at which geometrical theorems have application and physical laws are mathematical in character. It is, however, beyond my competence to try to flesh out or support this generalisation of the thesis, and it is not necessary to do so in order to solve the induction problem.

This brings us to the second point needing to be made. Although there is no relation of partial entailment binding probability evidence to conclusion, so that inductive or factual reasoning, intimately linked to probability, cannot be stamped with the authority of logic, it is still possible to argue, as in subsequent chapters, that our epistemic relations with the world form a system. Within that system such reasoning holds sway at all levels, including that at which transitive relations are picked out and lend themselves to inequality inferences and

those involving spatial and other properties. Thus, if I infer from two connected inequalities to a third, as in the example of the planks, which depends on such objects having a fixed size, then I may justify my belief about an inequality not yet observed by appealing to induction. The argument would be that I know from experience that planks are of fixed size, that directly comparing two planks, and then one of them with a third, does not cause them to expand or contract, there being no such effect, either, when I look at the planks and then close my eyes and look away. This illustrates a general truth, which is that there is more to induction than consciously noting particular regularities and then expecting them to continue, one's expectation being justified and caused by these past events and one's awareness of them. We appeal to induction, however vaguely, when we cite experience in justifying beliefs which, in a sense, we just have, as when I look at a plank and then close my eyes for a moment and expect it to be there when I open them. I have just tried to show, indeed, that such is its all-pervasiveness that it enters into the fundamental epistemic relations which give structure to our knowledge of the physical world.

These observations are meant to point the way to a solution to the problem of induction, which will gradually emerge from further consideration of other people's attempts, the next in line being the contribution made by R.A.Fisher, whose views will be discussed in the next chapter.

Chapter 2

Inverse Probability

It may be recalled that in the first chapter mention was made of inverse probability in the form of the no-miracles argument applied to sampling without replacement. On that occasion the reference was to the probability of the single ball left in a bag being black if the previous ninety-nine were, on which basis, or so I suggested, the probability would be very high. If it is now asked how our expectation as to the colour of the last ball is to be justified, then appeal may be made to the no-miracles argument, which states that in certain circumstances if H_1 and H_2 are competing hypotheses and if event e is more likely on H_1 than on H_2, then H_1 is more likely to be true, given that e has occurred. Suppose, then, that a bag is known to contain 100 balls, black or white. I draw 99 without replacement and they all turn out to be black, whereupon I reason as follows: on the hypothesis of 100 black balls originally in the bag, it is certain that the first 99 will be black; on the hypothesis of 99 black balls and a white ball, the probability of the first 99 balls being black is 1 in 100, which is very low. Therefore, it is likely that there were originally 100 black balls in the bag, which is to say that the remaining ball is black.

Now, there is a prima facie case to be made for the no-miracles argument being instrumental in deriving probability values from statistical data, for instance on the basis of a series of trials with a coin. If a coin about which nothing is initially known with regard to bias is tossed 100 times, always landing heads, so that the hypothesis considered is that $P(h) = 1$, and if the null hypothesis against which it competes is that the coin is a fair coin, then the following argument may be used. On the hypothesis that $P(h) = 1$ this is also the probability of 100 heads; whereas on the null hypothesis the probability is $(0.5)^{100}$, which is extremely low. But the outcome of the series of trials *is* 100 heads; therefore, $P(h) = 1$ is more likely than $P(h) = 0.5$. It does not follow, of course, that $P(h) = 1$ should be accepted, for the same argument would also favour any probability of heads greater than 0.5. It is, however, suggestive, and what it indicates is that a suitable form of the no-miracles argument could function as a sieve for possible hypotheses, provided that rejection criteria were decided in

advance, thereby creating a shortlist of candidates for acceptance. This, as we shall see later, is the approach adopted by R.A.Fisher.

Such applications of probability theory are, however, to be contrasted with the use of Bayes' theorem, for which in the present case we have the hypotheses

$H_1 : P(h) = 1; H_0 : P(h) = 0.5;$ and the evidence e: 100 tosses, all heads. If the theorem is

written as $P(H_1/e) = \dfrac{P(e/H_1) \times P(H_1)}{P(e/H_1) \times P(H_1) + P(e/H_0) \times P(H_0)}$,

then we have the following:

$$P(H_1/e) = \frac{1 \times P(H_1)}{1 \times P(H_1) + (0.5)^{100} \times P(H_0)}.$$

The problem here, as is well-known in such cases, is that of measuring the initial probabilities $P(H_1)$ and $P(H_0)$. We could equate them, in which case they cancel out to leave a very large value, almost 100%, for $P(H_1/e)$; but what does it *mean*?

This question can also be asked about the no-miracles argument, notwithstanding the fact that it enters into a great deal of our everyday reasoning or provides the grounds for it if required. This is a service it also performs for predictive scientific theories, for if such a theory is false and the predictions it generates are true, then in some cases it may be that agreement between observation and theory has to be explained away as due to chance, this being the null hypothesis, the rejection of which may imply acceptance of the theory. What the no-miracles argument does not do, however, is to exhibit a standard connection with frequency; although the probabilities of event e relative to particular hypotheses may themselves be thus connected, this is not true of the inference by which we pass from those probabilities to the conclusion that some of the hypotheses are more likely to be true than others. If in the case of one hundred balls in a bag, ninety-nine drawn and found to be black, I think there is a high probability that the last ball will be black, then I cannot ground my belief in a straightforward appeal to frequency, such that in most cases the last ball turns out to be black if all the others were, a concomitant of which is that there is no agreed method, either, of arriving at a numerical measure for the probability. If Laplace's rule of succession is rejected, and if it is indeed rational to believe that probably the last ball will be black, then perhaps the most I can say is that the probability increases with the number of black balls drawn, as would be the case if the bag contained one thousand balls, nine hundred and ninety-nine drawn and found to be black. In actual practice, of course, there are statistical methods

for extrapolating from a sample to a population; and these apply whether it is finite, as in the present case, or indefinitely large, as with tossing a coin.

It is worth noting that the no-miracles argument does not depend on there being a large number of positive instances of a property, as there are with ninety-nine black balls and the prediction that the last ball is also black. If it did, then perhaps the suspicion would arise that the argument was just a complicated form of inductive inference. As things are, it is easy to cite cases not involving numerous instances, as when all that is known is that a box contains several bags, which themselves contain either ninety-nine black balls and a white ball or ninety-nine white balls and a black ball. If a single ball is drawn, rather than numerous balls, and if it turns out to be black, then we have $P(b/b_{99}w_1) = \frac{99}{100}$, whereas $P(b/w_{99}b_1) = \frac{1}{100}$. It then follows, if the no-miracles argument is to be believed, that very likely the bag is of the first kind. What we still have to ask, however, is whether the argument can be justified.

Intrinsic probability relations and induction

One answer to this question is that the argument incorporates the notion of intrinsic probability, such that the relation between premiss and conclusion is again one of partial entailment, but this time not based on the principle of indifference. If this is to say that the argument employs *a priori* probability reasoning, and if it may be used to infer from observational premisses to unobserved events by way of probabilities, as with the drawing of the last ball, then clearly it poses a threat to scepticism about induction. The fact is, however, that theorists disagree about the analysis and justification of the no-miracles argument and inverse probability in general, especially with regard to the prickly question of initial probabilities; but then, it is also true that in practice there are well-established methods of sampling, or of deriving probability values from empirical data, without which there could be no work for statisticians and only a much restricted application of the probability calculus. What I have in mind, then, is to join the fray, one of my aims being to show that in fact the no-miracles argument does engage with frequency, and in such a way as to indicate that it operates within a system.

In his book *The Design of Experiments* (1935) R.A.Fisher, the statistician and probability theorist, attempts to anchor induction in hypothetic inference by considering cases in which a hypothesis leads to successful predictions. He discusses the example of a woman who claims to be able to discriminate by taste between tea in which milk has been poured first and tea in which it has been poured afterwards. To test this claim an experiment is

designed in which four cups with milk first and four with tea first are set out in random order, the woman having to place each in its category by tasting it. If she succeeds each time, and if we assume the null hypothesis that her claim is false, then it would seem that her guesses were correct by chance, the probability of which is $1 \div \dfrac{8!}{4! \times 4!} = \dfrac{1}{70}$. Since this is a very low probability, or so the argument goes, it is reasonable to discount the null hypothesis in favour of the woman's claim being correct.

It is possible, of course, that the proportion of correct guesses will be less than 100% but greater than expected as random choice, and that is why, according to Fisher, we have to decide in advance on what is to count as an acceptable significance level. Suppose, for instance, that the woman correctly places 3 out of 4 cups in each category, the probability of which is $\dfrac{16}{70}$ on the null hypothesis. If the significance level is set at 5%, which is a standard setting, then this means that a success rate is to be regarded as favouring the woman's claim only if the null hypothesis probability of its being a chance result is less than 5%. This rules out acceptance of her claim on the basis of 3 out of 4 correct answers, since $\dfrac{16}{70}$ is greater than 5%.

As Fisher points out, one advantage of such experiments is that they are open-ended in the sense that significance levels can be made as low as one pleases by increasing the number of cups. With 8 cups the probability of a 100% success rate is $\dfrac{1}{70} \cong 1.4\%$, so if the significance level is set at lower than 1.4% then the woman cannot establish her claim. But with 12 cups, 6 in each category, that probability is $\dfrac{1}{924} \cong 0.1\%$, and so on.

Given that such probabilities can be made as small as we please by extending the experiment, it might be expected that Fisher would argue for the possibility of decisively establishing the woman's claim; indeed, he does allow that she would have made good her claim, as he puts it (p.16) by achieving a 100% success rate. However, it would seem that he takes a falsificationist stance on confirmation, which leads him to say:

It might be argued that if an experiment can disprove the hypothesis that the subject possesses no sensory discrimination between two different sorts of object, it must therefore be able to prove the opposite hypothesis, that she can make some such discrimination. But this last hypothesis, however reasonable or true it may be, is ineligible as a null hypothesis to be tested by experiment, because it is inexact. If it were asserted that the subject would never be wrong in her judgements we should again have an exact hypothesis, and it is easy to see

that this hypothesis could be disproved by a single failure, but could never be proved by any finite amount of experimentation. (p.19)

There is a difficulty here, at least if one argues, as I did in the first chapter, that a great deal of certainty may attend particular factual beliefs, the evidence for which we may find compelling, as in the present case if we believe that the woman's correct answers cannot be explained away as a series of lucky guesses. Fisher speaks of proof, but what is in question here is not logical proof but the establishing of a factual claim by scientific method. In the present case it is a particular claim about sensory ability, but even if it were much vaguer, perhaps that the woman has *some* method of achieving her results, possibly by cheating, then it would seem that the inverse probability argument could still have a part to play.

Fisher himself sees a difficulty here, however, for if the argument depends on the improbability of a particular event, perhaps with a one in a million chance of occurring, and if this is to be understood in terms of its very low frequency, which itself is to imply that although it rarely occurs it sometimes does, then how do we know that this is not one of those times? Having raised the question, he goes on:

The 'one chance in a million' will undoubtedly occur, with no less and no more than its appropriate frequency, however surprised we may be that it should occur to *us*. In order to assert that a natural phenomenon is experimentally demonstrable we need, not an isolated record, but a reliable method of procedure. (p.16)

The trouble in the present case is that the advantage of a reliable procedure over an isolated record would seem to be just that a properly conducted experiment may be repeated or extended. Since Fisher says nothing about varying the conditions in the case of the tea-tasting woman, all that he seems to have in mind is that there is scope for a series of trials to be increased, as when the experiment with eight cups is repeated, the effect being to double the size of the first experiment. Essentially, then, we have a single experiment in which, depending on the number of cups, the probability of a particular chance detection rate could be less than one in a million. If Fisher says that the one chance in a million will undoubtedly occur, he should also say that the one chance in any number, however large, will also occur.

Although this view of the frequency connection with probability is debatable, it does not seem to be central to Fisher's reasoning, so that we still have a question as to the validity and significance of the no-miracles argument. It will help us to answer it, or at least to focus on a particular aspect of it, if we consider the views of someone else who is interested in the Fisher experiment. In *Hume's Problem* Colin Howson rejects any solution based on the no-miracles argument, whose use in the tea-tasting experiment he takes issue with for several reasons, one of which concerns the difficulty of ruling out alternatives to the woman's claim

if the null hypothesis is considered. He gives a list of possible ways in which the woman might know the answers, for instance by cheating, and continues, 'By the rules of chance the chance of a 100 per cent success-rate in repeated experiments with the eight cups if the null hypothesis is true is the total chance of all these alternatives. And we are back again to trying to calculate something that does not seem calculable.' (p.51) This is obscure, but it seems to imply that probability values need to be assigned to all possible alternatives if the woman's claim is to be tested. In that case it cannot be correct, for we can work out probability values only within a framework of beliefs, held steady for that purpose, so that we are not required to measure the probability of their being true. If I am about to draw a ball from a bag, which by previous inspection I believe to contain two black balls and a white ball, then in order to assign a probability value to its being black I do not need to do the same to the possibility of my having mis-remembered the colour proportions of the balls. This is partly because such a requirement would be unfulfillable, since the attempt to comply with it would generate memories of its own; and it is partly because when I say that the probability of black is two-thirds I imply first that this is the case if the bag contains two black balls and a white ball, and second that I believe that it does. If my belief turns out to be incorrect, my memory being faulty or someone having changed the balls without my noticing, then it is open to me to protest that I would have been right about the probability if my belief had been correct.

Similarly, it is possible that the tea-tasting woman is cheating, but the experimenters are required not to measure the probability of fraud but only to take precautions against it so that the possibility may be discounted. The question, in any case, is whether if she is not cheating her claim can be established by appeal to the no-miracles argument. And as I pointed out earlier, we could always regard the experiment as testing not the particular claim but only the possibility that she has a reliable method of ascertaining whether the milk has been poured first or last.

What we need to ask, however, is whether the no-miracles argument is germane to the problem of induction, to which the answer, I think, is that it depends on the logical status of the argument's inferential structure. According to Fisher, the force of the inference derives not from frequency considerations but from the notion of events having intrinsic probability. It is to this notion that we appeal when contrasting the unlikelihood of a particular event on one hypothesis with its likelihood on another, which is therefore more likely to be true. If that is the claim, then we should be able to assess it by the examination of significant cases, a method adopted by Howson in his criticisms of the no-miracles argument as providing a solution to Hume's problem.

Howson invites us to consider a particular case, the gist of which is as follows: suppose that an urn contains 999 white balls and 1 black ball, which together with 50 of the white balls is marked with the number 1. Let '1' mean 'ball is marked with the number 1' and suppose that a ball is drawn from the urn. Then $P(1/w) = \dfrac{50}{999} \cong 5\%$, and $P(1/b) = 1.$ Now suppose that the ball is marked with a 1; then according to the no-miracles argument, says Howson, we should reject at the 5% significance level the hypothesis that the ball is white. In fact, he says, $P(w/1) = \frac{50}{51}$ and $P(b/1) = \dfrac{1}{51}$, since there are 51 balls marked with a 1, only one of which is black. (p. 53)

Now, a problem with this counter-example is that it is obvious that Fisher, along with everyone else, would agree that $P(w/1) = \dfrac{50}{51}$ and $P(b/1) = \dfrac{1}{51}.$ To see what this indicates, we need to compare the example and the Fisher experiment at several points. First of all, take the null hypothesis in the experiment to be, as before, that the woman's claim is false and that her answers are correct by chance. Then if this hypothesis is ruled out, it is for two interwoven reasons: the first is that there is no prior reason to accept it; for instance, it is not the case that we are in the later stages of an experiment in which hitherto the proportion of correct answers has been no better than chance. Secondly, the outcome of the experiment gives us a reason for rejection, namely that all or most of the woman's answers have been correct, contrary to what would be expected if her claim was false.

Let us now make a comparison with the Howson example. Here the null hypothesis is that a ball drawn from the urn is white; this is to be rejected only if the same two conditions are satisfied: first, that there is no prior evidence in its favour; second, that there is a reason to reject it, this being furnished by the evidence of the ball drawn from the urn. In fact, there is very strong prior evidence that a ball drawn from the urn will be white; namely, that the urn contains 999 white balls and 1 black ball. Also, there is no posterior reason to reject the hypothesis, since a ball turning out to be marked with the number 1 is in fact very strong evidence that the ball is white. Since it is clear that the two examples are dissimilar in crucial respects, we should ask how the one informs the other with regard to their interpretation.

The Howson example is, in fact, a standard case of conditional probability, one in which the calculus of chances operates over prior probability values. It is in the absence of those values that Fisher would regard the no-miracles argument as having application; but what the Howson example shows is that if that argument is valid, then it depends on more

than there being a number of hypotheses which probabilify an event to varying degrees. The question of prior probabilities is one that Fisher does consider, for instance in his book, *Statistical Methods and Scientific Inference*. Here he discusses what he refers to as Michell's hypothesis that a particular constellation of stars – the Pleiades – is non-randomly clustered in the sense of being causally connected, the null hypothesis being that they are randomly distributed. He points out that in the absence of certain initial probability values it is not possible, *pace* Laplace and others, to assign a probability value *a posteriori* to the hypothesis of random distribution. This rules out the use of Bayes' Theorem, familiar to us in the shape of the conditional probability formula. He thinks all the same that the hypothesis may be rejected on the basis of significance levels, the same as with the tea-tasting experiment, which also was conducted in the absence of initial probabilities for the null hypothesis. So far, then, we are on familiar ground.

Given that on the null hypothesis the chance of a star belonging to a cluster is 30 in a million,[8] he thinks that it would be rejected on the basis of any reasonable level of significance. He goes on:

The force with which such a conclusion is supported is logically that of the simple disjunction: *either* an exceptionally rare chance has occurred, *or* the theory of random distribution is not true.

In view of the efforts which have been made to force a frequency interpretation on to such a disjunction, it is to be noted that the mental reluctance to accept an event intrinsically improbable would still be felt if, for example, a *datum* were added to Michell's problem to the effect that it was a million to one *a priori* that the stars should be scattered at random. We need not consider what such a statement of probability *a priori* could possibly mean in our astronomical problem; all that is needed is that if this datum were introduced into the calculation, then, in view of the observations, a probability statement could be inferred *a posteriori* to the effect that the odds were 30 to 1 that the stars really had been scattered at random. The inherent improbability of what has been observed being observable on this view still remains in our minds, and no explanation has been given of it. It has been over-weighted, not neutralized, by the even greater supposed improbability of the universe chosen for examination being of the supposedly exceptional kind in which the stars are *not* distributed at random. (p.39)

To elucidate the sense of this passage and to link it to Howson's criticism of the no-miracles argument, we would benefit, I think, from simpler examples than these writers have provided. Suppose again that a bag contains 100 balls, 99 drawn and found to be black. If the hypothesis under test is that all the balls are black, so that the last is, then the null hypothesis is that there are 99 black balls and a non-black ball, say a white ball, so that the last is white. Let the first hypothesis be H_1 and the second H_0. Using the usual notation, we have

---

[8] It would be tedious to have to trek through the abstruse mathematics of this example; in any case, I am about to give a very much simpler example of my own.

$P(b_{99}/H_0) = \frac{1}{100}$ and $P(b_{99}/H_1) = 1$. Fisher would now say that H_0 is to be rejected on any reasonable significance test; for instance, one that is based on a 5% or even a 1% level. The ground for this rejection, moreover, is that on the null hypothesis it is intrinsically improbable that the last ball to be drawn should be the white ball.

For Howson's criticism to be brought into play, we now have to modify this example in order to include prior probabilities. Suppose that the bag is one of 200 bags in a box, 199 containing 99 black balls and a white ball, the odd one out containing 100 black balls. Armed with these initial values we can now use the conditional probability formula[9] and thereby obtain: $P(H_0/b_{99}) = \dfrac{199}{299}$. This is to say that although $P(b_{99}/H_0) = \frac{1}{100}$ and $P(b_{99}/H_1) = 1$, so that the no-miracles argument would favour H_1, the initial values make it about twice as likely that H_0 is the case. Fisher would now maintain that the inherent improbability of the null hypothesis has been 'over-weighted' by the introduction of these values.

What this indicates is that if the point of the Howson example is to expose the invalidity of the no-miracles argument, then a possible rejoinder is that the argument derives its force from the intrinsic probability of events, even in the absence of initial values. If these are present, then this permits the application of Bayes' theorem, the effect of which is to yield posterior probability values, which again is an exercise in the intuiting of intrinsic probability relations. This is reminiscent of Keynes' thesis that a relation of partial entailment links probability evidence and conclusion.

Now, it is undeniable that in some sense probability is relative to evidence, and I shall argue later that this is a universal condition which relates evidence and belief within an explanatory framework. The point to be made here is that probability is Janus-faced, turning outwards to objective fact and inwards to rational belief. Insofar as it is relative to evidence, there seems little harm in qualifying probability as being intrinsic, if all that is meant is that it is always possible to appeal to evidence in favour of a probability judgement, or that the appeal is ultimate, perhaps in the sense that in defence of it all that we can do, apart from re-affirming the judgement, is to appeal to other evidence. What Fisher seems to have in mind, however, is that probability is intrinsic in a sense which disengages it from frequency, a consequence of which is that the no-miracles argument is valid even in the absence of prior probability values. If these exist and are known, so that Bayes' theorem may be used, then

---

[9] $P(H_0) = \frac{99}{100}$ and $P(H_1) = \frac{1}{100}$; so $P(H_0/b_{99}) = \dfrac{P(H_0 \cap b_{99})}{P(b_{99})} = \dfrac{\frac{199}{200} \times \frac{1}{100}}{\frac{199}{200} \times \frac{1}{100} + \frac{1}{200} \times 1} = \dfrac{199}{299}$.

instead of the no-miracles argument, with its focus on rejection criteria, we have a standard frequency-linked application of the probability calculus, as with the example just given in which $P(H_0/b_{99}) = \dfrac{199}{299}$. It is part of established theory that in such cases probability values have their empirical counterpart in frequency ratios. In the present instance this means that in a long series of trials in which a bag is chosen at random and ninety-nine balls drawn, the last ball turns out to be white in about two-thirds of cases in which the ninety-nine balls all turn out to be black. Alternatively, it means that in an indefinitely long series of trials the frequency approaches two-thirds as a limit.

What Fisher seems to believe, however, is that the thesis of intrinsic probability is opposed to a frequency interpretation of the no-miracles argument and other forms of probability reasoning, a point of view which accords with the everyday fact that the passage from evidence to conclusion need not be mediated by considerations of frequency. Thus, I may note the proportion of black balls in a bag and immediately pronounce this to be the probability of a black ball being drawn, all without any reference to frequency ratios in a series of trials. But then, it is true of any concept that we may apply it without being fully aware of its ramifications, as may be shown by consideration of the concept of colour. To begin with, my colour description means the same to me as to others only if certain conditions are fulfilled, for which a convenient shorthand is that we have a common system of reasoning and language use. Within that system public standards for the correct use of colour words translate into interactions between individuals. If I am one of those individuals, then the system operates not only over my similar colour experiences on different occasions, together with my being able to distinguish between the real and apparent colours of objects, but also over the colour perception and colour language behaviour of other people as it appears to me. That behaviour as observed by me is, of course, very different from my own colour perceptions, themselves very different from the look and sound of the language with which I refer to them. My experience of red is very different from my saying 'red' and from the appearance of other people when they observe and describe red objects; and yet, all these disparate elements combine within a system in which sounds and shapes become words. So much, then, for any use of the word 'intrinsic' in connection with colour descriptions.

Direct evidence and frequency

 In a similar vein, when I arrive at a probability value in a particular case I may reason from direct evidence. It is *because* I know that a bag contains 99 black balls and a white ball that I

believe that $P(b) = 99\%$ and expect a ball drawn from the bag to be black. But this is a matter of direct or immediate evidence, not of intrinsic probability. My beliefs make sense only within a system, and it is a fact about that system that probability entails frequency. What is true of standard probability is also true of its inverse counterpart. Yes, it is directly on the basis of 99 black balls being drawn that I expect the last ball to be black; or, if I am conversant with the no-miracles argument, then it may be directly on the basis of applying it in this case that I arrive at the same expectation. But it does not follow that the notion of intrinsic probability thereby commends itself, and in my view any insight to be gained from it is transferable to the notion of a system, which is more instructive and recognises the role of frequency.

If the intrinsic probability thesis is thus dismissed, then on the face of it the challenge to Hume must also lapse, for nothing in the notion of a system would seem to offer a solution to the problem of induction. This, however, may change as that notion is developed, to which end we have various matters to attend to, one of which concerns the frequency component of the no-miracles argument. As a first step, let us ask again what it is exactly that is problematic about inverse probability in the absence of initial values. Returning to our stock example of one hundred balls, ninety-nine drawn and found to be black, it just is a fact that most people would confidently expect the last ball to be black, the others falling into line if, say, the number of balls was increased to a thousand, nine hundred and ninety-nine of which had been drawn and found to be black. We are seeking to account for this fact by appeal to the no-miracles argument: $P(b_{99}/b_{99}w_1) = \frac{1}{100}$ and $P(b_{99}/b_{100}) = 1$; therefore, the last ball is likely to be black. But if that is the argument, then it seems to contain hidden premisses, the nature of which we are trying to uncover. It is a fact that we feel very confident that the last ball is black; given, however, that all thoughts of intrinsic improbability have been banished, we are left with a question as to how this confidence gains its credentials: if we draw ninety-nine black balls from a bag, do we really have the right to be sure that the remaining ball is black? It has to be said, after all, that the greater our certainty about the colour of the last ball, the greater our obligation to provide grounds and give reasons.

But what could these reasons be? Not intrinsic probability, nor any straightforward appeal to frequency. We are dealing, after all, with inverse inference from a sample to a population or from a sample to a probability; and all such samples are compatible with a subset of alternatives from a total set. In the present instance if ninety-nine black balls are drawn, then it would have to be maintained, in terms of a simple connection with frequency,

that in most such cases the last ball turns out to be black. In order to substantiate this claim we could ask someone to fill hundreds of bags with a hundred black or white balls, and then we could work our way through them, drawing ninety-nine balls from each bag and recording the colour of the last ball in each bag from which ninety-nine black balls have been drawn.

What could be simpler? Well, the drawback to this direct approach is that it allows a helper, if at all malicious, to fill each bag with ninety-nine black balls and a white ball. Suspecting that she may do this, we try to anticipate her by extending our claim to include the case in which ninety-eight black balls and a white ball, in any order, are drawn. We now assert that in both cases it is very likely that the last ball is black. Not to be outdone, she now fills every bag with ninety-eight black balls and two white balls. If our claim is limited to the two cases, we are not able to pronounce on the case in which ninety-seven black balls and two white balls are drawn, passing over them until we draw ninety-eight black balls and a white ball. We then announce that the last ball is black; unfortunately, it is white.

These examples illustrate a difficulty at the same time as they suggest a solution, for what is clear, at least with regard to balls in a bag, is that we need to consider systematic applications of the no-miracles argument to the full range of colour proportions for a fixed number of balls, not just the proportion in any particular case. The idea is that we take a complete set of alternatives and operate with significance levels for which the probability of success, as measured by the proportion of alternatives correctly rejected or accepted, is greater than is to be expected from blind guesswork.

Consider the example in which one ball is left in a bag after eight balls have been drawn. For reasons which will become clear, let the significance level be four ninths:

b_9; $P(b_8/b_9) = 1$; $P(b_8/b_8w_1) = \frac{1}{9}$; so $P(c) = 1$

This translates as follows: If the alternative is 9 black balls in a bag, the only possible sample is 8 black balls, the probability of which is 1; its probability given 8 black balls and a white ball is $\frac{1}{9}$; given this probability and a significance level of $\frac{4}{9}$ the possibility of 8 black balls and a white ball is correctly rejected, so that the probability of a correct rejection is 1. To continue:

b_8w_1; $P(b_8/b_8w_1) = \frac{1}{9}$; $P(b_8/b_9) = 1$.

$P(b_7w_1/b_8w_1) = \frac{8}{9}$; $P(b_7w_1/b_7w_2) = \frac{2}{9}$; so $P(c) = \frac{8}{9}$.

Translation: If the alternative is 8 black balls and a white ball, a possible sample is 8 black balls, with a probability of $\frac{1}{9}$; so the alternative is incorrectly rejected. Given the sample, the

other possible alternative is 9 black balls, given which the probability of the sample is 1; so this alternative is not rejected.

The other possible sample is 7 black balls and a white ball, with a probability of $\frac{8}{9}$; so $b_8 w_1$ is not rejected. Given the sample, the other possible alternative is $b_7 w_2$, given which the probability of the sample is $\frac{2}{9}$, so that the alternative is correctly rejected. So $P(c) = \frac{8}{9}$. To continue:

$b_7 w_2$; $P(b_7 w_1 / b_7 w_2) = \frac{2}{9}$; $P(b_7 w_1 / b_8 w_1) = \frac{8}{9}$

$\qquad P(b_6 w_2 / b_7 w_2) = \frac{7}{9}$; $P(b_6 w_2 / b_6 w_3) = \frac{3}{9}$; so $P(c) = \frac{7}{9}$

$b_6 w_3$; $P(b_6 w_2 / b_6 w_3) = \frac{3}{9}$; $P(b_6 w_2 / b_7 w_2) = \frac{7}{9}$

$\qquad P(b_5 w_3 / b_6 w_3) = \frac{6}{9}$; $P(b_5 w_3 / b_5 w_4) = \frac{4}{9}$; so $P(c) = \frac{6}{9}$

$b_5 w_4$; $P(b_5 w_3 / b_5 w_4) = \frac{4}{9}$; $P(b_5 w_3 / b_6 w_3) = \frac{6}{9}$

$\qquad P(b_4 w_4 / b_5 w_4) = \frac{5}{9}$; $P(b_4 w_4 / b_4 w_5) = \frac{5}{9}$; so $P(c) = 0$

$b_4 w_5$; $P(b_4 w_4 / b_4 w_5) = \frac{5}{9}$; $P(b_4 w_4 / b_5 w_4) = \frac{5}{9}$

$\qquad P(b_3 w_5 / b_4 w_5) = \frac{4}{9}$; $P(b_3 w_5 / b_3 w_6) = \frac{6}{9}$; so $P(c) = 0$

$b_3 w_6$; $P(b_3 w_5 / b_3 w_6) = \frac{6}{9}$; $P(b_3 w_5 / b_4 w_5) = \frac{4}{9}$

$\qquad P(b_2 w_6 / b_3 w_6) = \frac{3}{9}$; $P(b_2 w_6 / b_2 w_7) = \frac{7}{9}$; so $P(c) = \frac{6}{9}$.

And so on.

We see from the table that for most of the entries our confidence in rejecting alternatives is well served by the corresponding values of $p(c)$. Thus, if we draw eight balls from the bag and note that each of them is black, leaving b_9 and $b_8 w_1$ as the only possible alternatives, then we may reason as follows:

Given b_9, $P(b_8) = 1$, so $P(c) = 1$, since $b_8 w_1$ is correctly rejected.

Given $b_8 w_1$, $P(b_8) = \frac{1}{9}$ and if this is the sample then $b_8 w_1$ is incorrectly rejected. Also, $P(b_7 w_1) = \frac{8}{9}$, so $P(c) = \frac{8}{9}$, since $b_7 w_2$ is correctly rejected. This probability matches our confidence in rejecting one alternative and accepting another. Such reasoning is valid for eight out of the ten alternatives, the exceptions being $b_5 w_4$ and $b_4 w_5$, for which $p(c) = 0$. If there were 100 balls, 99 being drawn, there would be one exception, $b_{50} w_{50}$, with $p(c) = 0$, and for all the other alternatives the probability of success would be greater than 50%, which is the value associated with a random guess as to which alternative is correct. This would all

remain true if instead of using significance levels we simply rejected the alternative given which there was a lower probability of the sample than for the other alternative, which therefore is the one we would accept. Note, too, that the probability of success has a straightforward frequency connection; if, for instance, there are repeated trials in which eight balls are drawn from a bag containing eight black balls and a white ball, the cases in which they are all black will have a frequency of about one in nine, as will the incorrect rejection of the possibility that there are eight black balls and a white ball. But a sample consisting of seven black balls and a white ball will have a frequency of about eight in nine, as will the correct rejection of the possibility that the bag contains seven black balls and two white balls.

Let us now scrutinise these findings. First, it is clear that the malicious helper will fill the bag with five black balls and four white balls, or vice-versa, in which case we will always be wrong in the alternatives we reject. But then, of course, we will learn from the experience and bring our rejection criteria into line with it. Besides, there would not then be cases in which the sample was all black, or all white, these being the ones that interest us the most; and what this shows is that the device of the malicious helper is of limited utility. It just is a fact that before drawing from the bag we may have nothing to choose between the alternatives, this being a cognitive state of epistemic indifference which we often inhabit when faced with different possibilities.

A second point, which would seem to militate against the intrinsic probability thesis, is that if each sample is compatible with two alternatives, then there is nothing in the sample itself that could favour one of them against the other; it has no memory, as it were, of the colour of the ball left behind in the bag. And yet, the particular colour proportion of the sample does not just determine which two alternatives are in contention: it also probabilifies one of them at the expense of the other. It is *because* I have drawn eight black balls, or ninety-nine if the bag contains a hundred balls, or nine hundred and ninety-nine if it contains a thousand, that I reject one alternative and accept the other, thereby concluding that the last ball is black; and it is likely that I do so with a great deal of confidence which I feel to be justified.

If there seems to be tension between concepts here, then perhaps the way to mediate between them is to show that in other contexts they quite happily co-exist. We know, for example, that if a coin is repeatedly tossed, then the resulting frequency ratio, which we expect to continue, is consistent with a range of probability values for heads and tails, all of which give *some* chance of the frequency occurring. And yet, we take that frequency to favour some values over others.

Underdetermination and the notion of a system

It is arguable, even, that the marrying up of exactly the same evidence with a selection from a range of suitors in the form of possible hypotheses is, at some level, a cognitive process which is fundamental. Likewise with much of what we take for granted about the world, the evidence for which, in any particular case, is logically consistent with a wider range of possibilities. It is a cognitive process, then, which is exhibited in all forms of empirical inference, these being such that evidence and conclusion are only contingently connected. Thus it is that I may reach for a romantic novel from a bookshelf only to discover that what I thought was a terrace of books is in fact an empty façade, perhaps a prop from a stage play about suppressed passions in a library.

This example of sense deception leads naturally to the view that the link between similar evidence and different states of affairs extends even to the level of memory and recognition. If it seems to me that I recognise an object, say a table, then it is logically possible that I am deceived, as with the dummy book; and this may also be the case at a deeper level, that at which memory and expectation sustain from one moment to the next my belief that this is a table. Between the contents of my visual field a moment ago and my memory of them there is only a contingent connection, as also between my expectations and what will actually be seen a moment from now. And yet, it is *because* of my present perceptions of the table that I expect them to continue or to change in particular ways. If it is a table that I am looking at, then I do not expect to see a chair if I close my eyes and immediately open them.

It is true that to take this approach is to set aside considerations of probability; but the fact is that probability comes into play only if the inferential links between one moment and the next are pre-supposed. If I say that it is very unlikely that this table will suddenly vanish if I blink, then indeed it is but it depends on my being able to refer to a table as a distinct physical entity, itself possible only if my belief that there is a table is conveyed from one moment to the next by memory and expectation. Such a current of thought need only be diverted a little to lead to Hume's problem; but my point is that inferential linking at this deeper level is a fundamental cognitive process. It is essential to the flow of sensory information through a conceptual system which depends on it but also gives it form and content, without which there would just be a succession of unrelated moments.

There are themes here which we need to develop, but my present concern is with the ways in which they are relevant to our discussion of the frequency component of the no-miracles argument within a system of application. The example given earlier was that of nine balls in a bag, eight of them drawn, in consideration of which a systematic approach was taken to the inference from sample to population. By means of that approach I tried to show that there is a frequency connection with the no-miracles argument, one which holds in at least some cases of inverse probability in the absence of initial probability values. What I now propose is to continue to investigate probability as a system, one that we may then place in the wider framework of general inductive reasoning.

To begin with, let us enquire into the statistical methods by which the probability of heads is estimated from a series of trials in which a coin is tossed. These methods, or at least their interpretation, may vary according to the school of thought, and it is easy to imagine what Fisher's approach would be. Given a particular outcome, a sequence of heads and tails, the probabilities of that outcome for different values of $P(h)$ would be compared, perhaps with reference to a test hypothesis, a favoured value of $P(h)$. A basic principle is that after numerous trials $P(h)$ is taken to be the same as the proportion of heads, provided that heads and tails are distributed according to certain conditions. Relevant here is the fact that the conditional probability of that proportion, given a probability of heads, is at its maximum when proportion and probability of heads coincide. Suppose, for instance, that the result of ten trials is given by $h_9 t_1$, and let $P(h) = p$; then $P[h_9 t_1 / P(h) = p] = p^9 (1 - p)^1 \times {}^{10}C_9$. That is, the probability of $h_9 t_1$, given that $P(h) = p$, is $p^9 (1 - p)^1 \times {}^{10}C_9$, which reduces to $10p^9 - 10p^{10}$. By means of the differential calculus, it can be shown that this is at a maximum when $p = \frac{9}{10}$ for $0 < p < 1$.

If $P(h)$ is greatest when equated with the proportion of heads, then clearly this maximum will change at the next toss of the coin, in tandem with a change in the proportion of heads. What this indicates is that if the experiment continues then certain tests have to be performed before a particular value of $P(h)$ emerges as the clear favourite. Whatever that value, and however strongly confirmed, we always have to allow for a range of deviation with regard to frequency. Thus, if we toss the coin 1000 times, having decided along the way that $P(h) = \frac{9}{10}$, and if the overall result is given by $h_{920} t_{80}$, we do not then announce that $P(h) = \frac{920}{1000}$; instead, we say that this deviation from the estimated value of $\frac{9}{10}$ is within

acceptable limits, given that the estimate is a best fit for the relative frequency and distribution pattern of heads and tails in the sequence.

This is to imply, of course, that some deviations are not acceptable, the line being drawn by the use of significance levels. It would then seem to follow that the no-miracles argument is required at the basic level of inferring between frequency and probability. To this, however, it might be objected that we may appeal instead to a standard frequency connection, for instance with regard to $P(h) = \frac{9}{10}$. Thus, If you take thousands of coins and toss them a thousand times, selecting those that yield an acceptable proportion of heads and tails for a given significance level, and if you keep tossing these coins, then you will find that for most of them the probability of heads turns out to be nine-tenths, that for very few of them does it turn out to be half, and that for almost none of them does it turn out to be very low.

If that is the argument, then, as before, it leaves out of account the possibility of our malicious helper ensuring that all the coins are biased towards tails, so that none of the coins we select will have a probability of heads of nine-tenths, or even a probability greater than half. And the only way to overcome this difficulty, again as before, is to locate a particular inference within a general system, in this case a system in which any outcome of a series of trials will form the basis for a probability estimate.

Very well, but suppose we appeal to frequency in a series of trials using just this coin, not thousands of coins. The difficulty here is that whatever the sequence of heads and tails, it is consistent with their having any probability whatsoever, a consequence of which is that we have to justify our choice of candidate. Surely, then, there must be criteria by which an infinite range of possible values of $P(h)$ is rejected, a result achieved by appealing to the improbability of the sequence, or of the relative frequency of heads therein, given these values of $P(h)$. To speak in this way is to appeal to the notion of a system with regard to inverse probability, as in the case of the nine balls in a bag, so that again the no-miracles argument is brought into play; it has, too, a frequency component in the form of the success rate of the inferences we make, in this case from sequences to the probabilities of heads and tails. With regard to that success rate, it would be much too tedious to have to go into the kind of detail by which a systematic approach was shown to be necessary in the case of nine balls in a bag. What we can say, as a labour-saving device, is that statistical methods have met with a great deal of success, even if there is not always a consensus as to how they are to be applied or understood.

Let us now return to our earlier rebuttal of Howson's criticisms of the no-miracles argument. In his discussion of the tea-tasting experiment, he considers alternatives to the woman's claim being correct, including the possibility of her cheating, and he maintains that in order to assess her claim we need to be able to calculate the probabilities of these alternatives. This, I have argued, is a mistake, for there has to be a sense in which probability is relative to the available evidence. Thus, when it comes to estimating probabilities from the tossing of a coin, we need not consider extrinsic factors such as the large but for practical purposes unquantifiable proportion of fair coins in circulation. Otherwise, we should not be able to apply criteria by which the fairness of coins was judged – or not if it depended on a prior probability assessment of other coins. We need to keep in mind that when we pick up a coin it may be regarded as one of a global set of millions of coins with differing probabilities of heads, nothing being known about the proportions of these probabilities, which are therefore regarded as being epistemically neutral. It is true that one may safely assume, or at least that one does assume, that most British coins are fair, even if those from other countries are suspect; but what this shows, apart from prejudice against foreign mints, is that in ascribing probability we disregard such facts and rely on internal evidence from a series of trials with the particular coin. This procedure would make no sense if probability criteria had to take account of such facts, for then they would have no application and it would not be possible to speak of a coin being fair or biased.

Finally, or before the summing up, let us probe the semantic content of reference to frequency in a probability context. What does it mean to speak of the frequency of heads and tails and how is it connected with rational choice? Relevant here is the fact that the raw materials of frequency need to be interpreted within a probability system. If heads predominate when I toss a coin several times, then I do not need the calculus of chances to work out that heads is more likely than tails, or at least that it would be rational to entertain that belief, for this is frequency only in the sense of one face landing more often than the other. I do need the calculus, however, if my concern is with the frequency of heads as a stable, numerical property of the coin, or of tosses of the coin, one that emerges in more refined detail as the number of trials increases. It is frequency thus characterised which in some sense remains the same; it allows for the fact that the proportions of heads and tails in any sequence will change at each new toss of the coin. Clearly, this more sophisticated concept, that of statistical frequency, belongs within a probability system.

But then, this must also be true of the no-miracles argument, for it engages with conditional probability relations between evidence and hypothesis or sample and population.

These involve frequency, and the case of nine balls in a bag shows that there is also a frequency connection with the probability of successful inference. This all depends on the extraction of probability values and frequency ratios from given data, for instance in the form of a particular sequence of heads and tails. Crucial to this process is the sieving out of possibilities by inverse probability methods, which essentially are those of the no-miracles argument. It follows that this argument depends on frequency and probability, which themselves depend on the argument.

And now for the summing up. If it can be shown that the no-miracles argument is an essential tool of science and of everyday thought, including the inductive inference by which, or so it seems, experience generates knowledge, then doors to the solution of Hume's problem do not thereby automatically open. Hume, after all, makes it very difficult to gain admittance, and he need only point out that the argument, like all probability arguments, ultimately presupposes the validity of inductive reasoning. We shall return to this, the present point being that the argument is thought by Keynes and others to hold a key to a solution if accompanied by a thesis of intrinsic probability. That thesis, in my opinion, is improperly dressed and needs to be fitted out in the language by which the notion of a system is expressed. This is to say that Fisher's views, as also those of Keynes, are at least knocking at the right door, for there must be entailment in *some* sense, otherwise there can be no solution. They are blocked, however, by too narrow a view of the sceptical argument, one inherited from Hume, and fail to see that it applies to the premisses of an inductive inference just as much as to the passage to a conclusion. This will be explained later, having only been hinted at in the present chapter, which has been more concerned to show that probability reasoning forms a system; the possibility then arises that the same may be true of inductive inference in general, with which it is interwoven.

We have much to do before the doors open, one key to which is further consideration of attempted probability solutions to the problem, which is how I propose to occupy myself in the next chapter. In the meantime, perhaps I could end this one by illustrating the no-miracles argument as it appears in everyday reasoning. Suppose that I call in at my friend's haunted castle every day and that I have never known her to fail to answer the door when in residence. Today, however, I bang the knocker to no avail, its echoes in the hall not portending the clank of chains as she descends the stairs. I now reason as follows: on the assumption that she is in, the probability of her not answering the door is very low, especially in view of the fact that because she is a ghost there is very little to delay her in the way of being tied up on the phone or in the bathroom or lying dead in bed. But if, on the other hand, she is out, then her not

answering the door is a certainty; therefore, she is likely to be out. This, surely, is a form of probability reasoning that we often employ, the legitimacy of which it would not occur to us to question. In the present case, as in that of the nine balls in a bag, a significant point, given that such reasoning belongs within a system, is that the workings of that system are *hidden*.

Chapter 3

Stove's Attempts at a Solution

First attempt

My aim in this chapter is to finish interviewing, as it were, the probability candidates for a solution to the problem of induction, and to deliberate on whether this last candidate, the thesis presented by the Australian philosopher D.C.Stove, succeeds in its arguments. Having scrutinised that thesis, I shall then be concerned to field my own candidate by developing the view that inductive scepticism cannot stop short of wholesale rejection of empirical knowledge, a position it can occupy only by refuting itself.

Inductive scepticism, according to Stove, is universal in scope and can therefore be defeated by the production of a single counter-example. The one he has in mind involves a particular question in ornithology: can it be proved that unobserved ravens are likely to be black if most observed ravens have been black? He claims that it can, and his method is to appeal to logical probability, one difficulty with which, he says, is that the reader may remain unconvinced by any demonstration based on manipulating symbols, even if he or she is unable to pinpoint any particular weakness in the chain of reasoning.

Such reactions against logical probability arguments are remarked upon in *The Rationality of Induction*, in which he attempts to substantiate his claim about the blackness of ravens. The sceptic targeted by Stove is David Hume, who, he believes, was at one time sympathetic to what he referred to as scepticism with regard to the senses. Stove takes this to be scepticism about the premisses of inductive inference, since these are statements about the evidence of our senses. He thinks that this is different from scepticism about the inference to unobserved events and is independent of it. His concern, he says, is only with this latter kind of scepticism. The distinction made here is one that seems to me to be seriously wrong, but a discussion of this point will have to be deferred until we have examined his rejoinder to Hume.

His approach to the problem is to give a probabilistic interpretation to certain arguments of Hume's, and to this end he quotes from *An Enquiry concerning Human Understanding* the following passage in which Hume is discussing the possible effects when billiard balls collide. Prior to experience, Hume says, the striking of the cue ball against the target ball may be followed by any of a hundred different events:

May not both those balls remain at absolute rest? May not the first ball return in a straight line, or leap off from the second in any line or direction? All these

suppositions are consistent and conceivable. Why then should we give the preference to one which is no more consistent or conceivable than the rest? (1975, p.29)[1]

Stove now suggests that Hume's thesis is what may be called a judgment of irrelevance:

Schematically speaking, a judgement of irrelevance is the assertion that q is neither more nor less probable in relation to r-and-p than it is in relation to r alone. In short, p is irrelevant to q in relation to r, if and only if $P(q/r.p) = P(q/r)$. And Hume's sceptical thesis about induction was at least the thesis that, in relation to propositions knowable *a priori*, propositions about the observed are in this sense irrelevant to propositions about the unobserved. (Op. cit. p.38)

[Note that he uses a dot instead of the set notation intersection sign we used in previous chapters]

Citing evidence for this interpretation, he points out that Hume's contention was that we have no reason to believe any proposition about unobserved events *a priori*, or, 'in other words, in relation to propositions knowable *a priori*' (Ibid). This should give us pause, for it raises a question as to Stove's understanding of Hume, but let it pass for the moment. Keeping in mind, then, that he denotes these latter propositions by the letter r, we can guess what it is that he is trying to do: he wants to be able to write $P(q/r.p) = P(q/r)$ as a probability expression of Hume's argument. There are, however, difficulties here, and they are best approached by noting, first of all, that these conditional probability expressions belong within a formal system in which they are subject to rules, one of which yields the result that

$$P(q/r) = \frac{P(q.r)}{P(r)}.$$ Expressed in this way, as a probability fraction, $P(q/r)$ is usually

understood to be the numerical ratio of the conjunction of the sets of alternatives q and r to the set of alternatives r.[2] This ratio is then calculated according to the theory of combinations, except that nothing like that applies in the present case, where the point is not to calculate these expressions but to manipulate them according to the rules of logical probability.

Now Hume's original statement was that we have no reason to believe any proposition about unobserved events *a priori*, which Stove interprets in a particular way. So, if q and r have the same meaning as before, and if it is legitimate to write $P(q/r)$ with regard to the statement, as Stove thinks it is, then Hume should be understood as claiming that $P(q/r) = 0$, since what he says is that we have no reason to believe q given r, and I do not

---

[1] Quoted by Stove on page 37.
[2] This ratio may be illustrated by the use of Venn diagrams. See page 65.

see how else the negative part of the statement could be expressed. But for any q and any r, $P(q/r) + P(\sim q/r) = 1,$ and this reads: the probability of q given r plus the probability of not-q given r equals 1.[3] For q and r in their present meanings, where $P(q/r) = 0,$ it follows that $P(\sim q/r) = 1.$ But $P(r) = 1$, since r is an *a priori* proposition, therefore $P(\sim q) = 1.$

I need hardly point out that this is not a sustainable interpretation of Hume. In the billiard ball passage, for instance, he says that there is no *a priori* reason to believe that a particular event will occur. If q denotes that event, and if $P(\sim q) = 1,$ then Hume is claiming that the probability of the event not occurring is one hundred per cent. This is not a claim that could ever be made by a sceptic about induction.

As we saw in the passage quoted from Stove, he uses the expression $P(q/r.p)$, where p stands for observational premisses. He now claims that Hume's contention was that we have no reason to believe any proposition about the unobserved even when it is based on observational premisses. Accordingly, he says, we must 'suppose that Hume, when he said that the premiss of an inductive inference is *no reason* to believe the conclusion, meant at least that it is no *more* reason than a proposition knowable *a priori* is' (p.39). He thinks that this entitles him to write $P(q/r) = P(q/r.p),$ but, as I have just pointed out, it follows from Stove's own reasoning that $P(q/r) = 0$; therefore, $P(q/r.p) = 0$ and, as before, $P(\sim q) = 1.$

I think we are now in a position to object to this whole programme of trying to fit Hume into the procrustean bed of a formal probability system. With this in mind, consider the billiard ball passage again, in which Hume claims that we have no *a priori* reason to expect the target ball to behave in a particular way. We know from the *Treatise*, as Stove points out, that in separate passages Hume expresses the view that we have no reason even after repeated observation of the events following one ball striking another. It is these passages that Stove has in mind when he writes $P(q/r) = P(q/r.p).$ That being the case, we are free to let p denote the target ball being struck by the cue ball and q denote the target ball's trajectory according to physics. Now let us ask whether a probability interpretation in terms of p and q can be given to the passages about billiard balls. In those passages Hume says that the target ball could do any of a hundred things after being struck and that there is no reason to favour any particular one of them. When he talks about a hundred things he means an indefinitely large number, since no limit can be placed on the number of things imaginable. In that case, and even if a probability interpretation is appropriate, it is not clear whether we should write

---

[3] See the Venn diagrams on page 65.

$P(q/p) = 0$ or $P(q/p) \cong 0$, which reads: the probability of q given p is more or less equal to 0. It would have to be one or the other, so let it be the former.

We know that $P(q/p) = \dfrac{P(q.p)}{P(p)}$, and, as we have seen, this formula is used to calculate the numerical value of probability ratios. Since nothing like that applies here, the only way we can adapt the formula is by letting $P(p) = 1$. But this gives $P(q/p) = P(q.p)$, where there is no useful distinction between $p(q)$ and $P(q.p)$. We might as well incorporate p into the initial conditions in terms of which q is defined. But this empties the formula and leaves us with $P(q) = 0$.

A parallel case would be that in which there are 10 balls in a bag, 5 white and 5 black, with A denoting: 'the first ball is black', so that $P(A) = \frac{1}{2}$. But now, and to no purpose, it occurs to us to let B denote: '10 balls in a bag, 5 white and 5 black' in order to be able to write $P(A/B)$, except that we now have to let $P(B) = 1$.

Let us now try to consolidate these gains. If q stands for the target ball having a particular trajectory, then we may ask whether it makes sense to speak of the probability of q when nothing could count as confirming or disconfirming it. According to Hume, we have no reason to predict q, either before or after experience, and this means that even if q does occur, the event being repeated every time the target ball is struck in a particular way, there is still no reason to believe that q will occur on any future occasion. If q is predicted to occur and does occur, this is not to be regarded as confirming the prediction in the sense of legitimising it, nor is it to be taken as evidence for any future occurrence of q or, indeed, of any other event.

But now, if we are to speak of the probability of q, then what kind of probability could it be? To see the difficulty here, imagine a person saying that the probability of a particular coin turning up heads is half. Now he throws it a thousand times and on each occasion it turns up heads. If he still insists that the probability of heads is half, then perhaps the probability in question is regarded as being *a priori*. He may mean that with a fair coin the probability of heads is always half, this being true by definition; and perhaps it simply hasn't occurred to him to revise his estimate in the light of the new evidence.

A priori probability, or classical probability as Ayer calls it, has nothing to do with actual events, and that is why the probability of heads with a fair coin is always half. As Ayer puts it, to say that the probability is half, given that the coin has two faces, is to say no more than that one is half of two. (1972, p.28) Could this be the kind of probability symbolised by

$P(q) = 0$? If so, then the reference would be to q as one among an indefinitely large number of equal possibilities, and to say that $P(q) = 0$ would be to say that the ratio of one to very many is, or tends to, zero. This *a priori* interpretation would be consistent with Hume not allowing predicted outcomes to have any bearing on his assertion that we have no reason to believe *q*. All the same, it cannot be correct, since it would mean that Hume was asserting only that one to many is almost none, which, though unarguable, is no more than a truism.

Now that *a priori* probability has been ruled out, let us return to the example of tossing a coin a thousand times, the outcome always being heads. If the person insists that the coin is a fair coin, it may be that he bases his judgement on previous trials, the evidence from which outweighs the present run of heads. This, of course, is to recognise that such outcomes are relevant to the question of probability values. The sense of probability here is such that one's assessment of the probability of heads may be subject to revision after further tosses of the coin, in line with the fact that this concept of probability has empirical content as an essential component. None of this applies to the statement that $P(q) = 0$, since, if Hume is correct, it makes no difference how many times the target ball has behaved as expected. It follows that either $P(q) = 0$ makes no sense or it is an *a priori* probability statement, in which case it cannot be regarded as an interpretation of Hume.

There are difficulties in understanding Hume, and it may be that we can turn them to our advantage when it comes to answering his sceptical thesis. What needs to be stressed at the moment is that these difficulties work against Stove, whose proof about ravens depends on a probability interpretation of Hume. That being the case, it may be asked whether there is any point in going on to discuss his proof of a counter-example to Hume, since it consists in a chain of reasoning with $P(q/r) = P(q/r.p)$ as one of the main links. The point, I think, is that when we emerge from the other side of the discussion we shall find ourselves better placed to answer a key question about probability and induction: that of whether probability theory has any part to play in the solving of Hume's problem.

Perhaps we may start by examining Stove's proof that the blackness of observed ravens is relevant to the statement that unobserved ravens are black. His method of setting out his proof is to proceed by way of numbered statements with cross-references from one to another. I shall refer to his statements using the same numbers as him, but mine will not be consecutive, since I don't propose to comment on every statement he makes.

He begins by substituting *t* for *r*, where *r* denoted an *a priori* proposition and *t* now denotes a tautology. He claims that, 'Part at least, then, of the content of Hume's sceptical

thesis about induction is this: (14) for all e and all h such that the inference from e to h is inductive, and for all tautological t, $P(h/t.e) = P(h/t)$.' (p.40) He now displays his tools, as it were, in the form of the principles he intends to use. These are the negation principle: $P(q/r) + P(\sim q/r) = 1$ and the conjunction principle: $P(p.q/r) = P(p/r) \times P(q/r.p)$. (p. 48) What I now propose is to explain the first of these principles by the use of Venn diagrams, which may then be used to illustrate Stove's proof.

A Venn diagram is a set of interlinking circles used to display and calculate probabilities. The following diagram represents the probabilities of p and q.
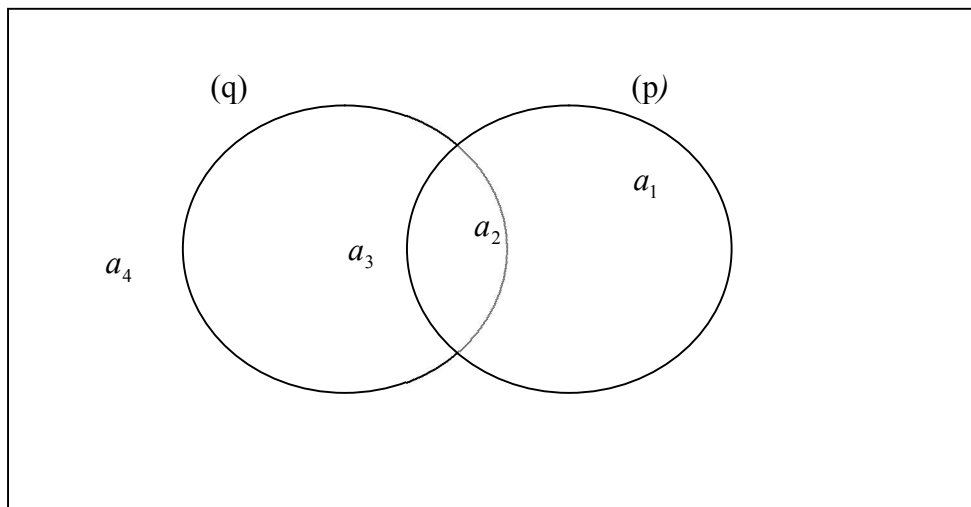
The rectangle represents the total possibility space, therefore $a_1 + a_2 + a_3 + a_4 = 1$. The overlap between circles p and q represents $P(p.q)$, which therefore $= a_2$. Similarly, the probability of p but not $q = a_1$, of q but not $p = a_3$; and the probability of neither p nor $q = a_4$.

$P(q/p)$ is taken to be the ratio of the conjunction of p and q to p; that is, $P(q/p) = \dfrac{a_2}{a_2 + a_1}$.

Similarly, $P(p/q) = \dfrac{a_2}{a_2 + a_3}$.

We are now able to comprehend the inner workings of the conditional probability formula $P(q/p) = \dfrac{P(q.p)}{P(p)} = \dfrac{P(p/q) \times P(q)}{P(p)}$:– $\qquad\qquad P(q/p) = \dfrac{a_2}{a_2 + a_1}$,

$P(q.p) = a_2$ and $P(p) = a_2 + a_1$; therefore $P(q/p) = \dfrac{P(q.p)}{P(p)}$. The same kind of thing can be

done for $P(q/p) = \dfrac{P(p/q) \times P(q)}{P(p)}$.

To demonstrate the negation principle $P(q/p) + P(\sim q/p) = 1$:–

$$P(q/p) = \frac{a_2}{a_2 + a_1} \quad \text{and} \quad P(\sim q/p) = \frac{a_1}{a_1 + a_2}; \text{ therefore } P(q/p) + P(\sim q/p) = 1.$$

To demonstrate the conjunction or multiplication principle $P(p.q/r) = P(p/r) \times P(q/r.p)$, we should have to add another circle to represent r, where that circle intersects the other two; but I think that we can manage without it.

We now come to Stove's attempt to prove that there is an evidential connection between observed and unobserved ravens. His approach is to try to prove the falsity of statement (14): For all e and all h such that the inference from e to h is inductive, and for all tautological t, $P(h/t.e) = P(h/t)$. He starts by using the conjunction principle to derive statement (36): $P(p/t) \times P(q/t.p) = P(q/t) \times P(p/t.q)$. From (36) he derives (37): If $P(p/t)$ and $P(q/t)$ are each >0, then $P(q/t.p) = P(q/t)$ if and only if $P(p/t.q) = P(p/t)$.[4] (p.51) Interpreting (37) in terms of our diagram, we have:

$P(q/t.p) \ = \ P(q/t) \quad$ If and only if $P(p/t.q) \ = \ P(p/t)$.

$\dfrac{a_2}{a_1 + a_2} \ = \ a_2 + a_3 \quad$ If and only if $\dfrac{a_2}{a_2 + a_3} \ = \ a_2 + a_1$. Note that $P(t) = 1$ and is a

dummy constant in this manipulating of terms.

So (37) turns out to be a mathematical truism about transposition of terms. If we were to work our way through the numbered statements in Stove's proof, we would find that each of them corresponds to an equation based on the Venn diagram and derived from transposing and collecting terms. Here, for instance, is statement (42) together with its associated equation:

$P(q/t.p) \ = \ P(q/t)$ if and only if $P(q/t.\sim p) \ = \ P(q/t)$. (p.52)

$\dfrac{a_2}{a_2 + a_1} \ = \ a_2 + a_3$ if and only if $\dfrac{a_3}{1 - (a_2 + a_1)} \ = \ a_2 + a_3$.

---

[4] The book has $P(q/t)$, but it should be $P(p/t)$.

The easiest way to show that this is correct is as follows: If the stated condition is

$$\frac{a_3}{1-(a_2+a_1)} = a_2 + a_3, \text{ then } \frac{a_3}{a_2+a_3} = 1 - (a_2+a_1). \text{ But } \frac{a_3}{a_2+a_3} = \frac{a_3+a_2-a_2}{a_2+a_3} = 1 - \frac{a_2}{a_2+a_3}$$

So the condition may be expressed as:

$$1 - \frac{a_2}{a_2+a_3} = 1 - (a_2+a_1), \text{ from which it follows that } \frac{a_2}{a_2+a_1} = a_2 + a_3, \text{ as required.}$$

In statement (42) Stove now substitutes H for q, where H denotes: 'there are Australian ravens and all of them are black'. For p he substitutes E: 'All the observed Australian ravens have been black'. Statement (42) now becomes statement (47): $P(H/t.E) = P(H/t)$ if and only if $P(H/t.\sim E) = P(H/t)$. It follows immediately that $P(H/t.\sim E) \neq P(H/t)$, and this is the conclusion of Stove's proof. It follows because $P(H/t) > 0$ but $P(H/t.\sim E) = 0$, since it is not possible that all ravens are black if not all observed ravens have been black. This, of course, is a trivial analytic truth, but it is all that Stove's proof amounts to. In order to render his proof invalid as a counter-example to inductive scepticism, all we need do is to substitute H' for H, where H' denotes 'All unobserved Australian ravens are black.' It cannot be objected to this procedure that we have altered the reference of H, since this is the same for H' and E as it is for H. But now the proof collapses, since it is not true that $P(H'/t.\sim E) = 0$.

Second attempt

I conclude that Stove's proof is invalid and does not establish a counter-example to Hume. Even if it were valid, its relevance to Hume's thesis would be questionable because of the inconsistencies that arise when his arguments are rendered into probability language. Stove does, however, have another proof in favour of inductive inference, and although it involves probability, it does not depend on a probability interpretation of Hume and must therefore be treated separately from the arguments we have just discussed. It is to this proof that we now turn.

His approach this time is to consider the possibility of inferring to a population on the basis of the properties exhibited by a large sample. Taking this as an inference from the observed to the unobserved, he thinks that it will serve as a counter-example to Hume if its validity can be established in any particular case. An inference of this kind is an example of the inverse probability we discussed in chapter one in connection with balls in a bag.

Stove is aware of the controversial nature of inverse probability in the absence of initial probability values. Such inferences can be justified, he thinks, not in terms of Laplace's mathematical arguments, which we have already discussed, but by an appeal to the principles of probability. The basic idea, according to him, is one that was formulated by D.C.Williams in *The Ground of Induction*. Stove thinks that the book's main argument is essentially correct, and his own proof of induction is a development of that argument. What he attempts to prove is that if 95% of a 3000-fold sample of ravens are black, then there is a probability greater than 90% that the proportion of black ravens in the population matches the sample to within 3%; i.e., that 92% - 98% of the population are black. He thinks that if this can be proved, then it constitutes a counter-example to Hume, a claim to which I shall return after examining his attempt at a proof.

In his preliminary remarks Stove asks us to consider a population of a million ravens and a 3000-fold sample and to compare them with regard to the proportion of black ravens in each. He defines an approximate match as one in which the proportion of black ravens in the sample is within 3% of the proportion in the population. He is now able to state, correctly, that the probability of an approximate match is least when the proportion of black ravens in the population is 50%. This is true of small as well as large numbers; for instance, if you consider 200 ravens, 100 of which are black, and if you take a sample of 100, then the sample matches to within 3% if 47%-53% of the sample is black; i.e. 47 to 53 of the ravens, for which the probability is 0.68. If you now calculate the corresponding probabilities for all the other cases, from no raven in the population is black to all 200 ravens are black, you will find that they are always > 0.68.

If a population of a million ravens contains 50% that are black, and if a 3000-fold sample is taken, then it is possible to work out the probability of a match within 3% and this turns out to be >0.9. This is the probability that 1410 – 1590 ravens will be black. Having stated this fact, Stove now makes what he regards as a key statement, hence the italics: '*Whatever the proportion of black ravens may be in a population of a million, at least nine out of ten 3000-fold samples of that population do not diverge from that proportion by more than 3 per cent in the proportion of black ravens they contain.*' (p.70) Again, this is correct: if the probability is > 0.9, then it follows that at least nine out of ten 3000-fold samples are favourable ones. This is true whatever the proportion of black ravens in the population, since the probability of an approximate match is least when that proportion is 50%, and then it is still >0.9.

Stove now makes an assertion which seems, on the face of it, to be counter-intuitive. He says that his key statement is true whatever the size of the population, provided, of course, that it has at least 3000 members. One would expect the probability of an approximate or exact match to decrease as the size of the population increases, and this indeed is what happens. The probability is 100% when the population is composed of 3000 ravens and thereafter decreases as the population increases from 3000. Nevertheless, Stove is correct, and for a reason he omits to mention; namely, that the probability approaches a finite limit as the population increases indefinitely.

To understand the way in which limits operate in this class of cases, we lose nothing by taking the simplest possible example, which is that in which a 2-fold sample is taken from an x-fold population in which 50% of the ravens are black. For our present purposes, we may allow the matching to be exact, and this occurs when one raven in the sample is black, the probability of which is $\dfrac{\dfrac{x}{2} \times \dfrac{x}{2} \times {}^{x^2}C_1}{x(x-1)}$. This $= \dfrac{x}{2x-2}$. Dividing numerator and denominator by x, $\dfrac{x}{2x-2} = \dfrac{1}{2 - \dfrac{2}{x}}$, which approaches $\frac{1}{2}$ as x approaches infinity. Since this limit is a minimum, it follows that in approaching $\frac{1}{2}$, the probability is always $> \frac{1}{2}$. In the example given by Stove the limit is >0.9 and is a minimum, so that the probability of a match within 3% is always >0.9 whatever the size of the population.

Before turning to Stove's proof, I think we need to be clear as to what it is that we agree with him about. Perhaps the best way to do this is to ask what use could be made of certain facts. Assuming a 3000-fold sample of ravens, it is a fact that the probability of a match within 3% is > 0.9 whatever the size of the parent population and whatever the proportion of black ravens it contains. I can use these facts if I know what the population proportion is, because then I can be 90% certain about the approximate proportion of black ravens in the sample. Also, these statements about probability can be translated into statements about the ratio of favourable 3000-fold samples to the total number of 3000-fold combinations extracted from a population. In other words, there is a standard frequency link. All this is clear, including the fact that any inference I make is from a population to a sample and can be made if I already know the proportion of black ravens in the population. Stove would have to agree, of course, that this is a factual, non-logical inference, at least at this stage of his argument, otherwise it would already constitute a counter-example to Hume. What he wishes to be able to prove, using logical probability, is that it is possible to start with

a sample and infer from the sample to the population, where this latter inference does have logical force.

Whether he succeeds in this, and what the significance for induction may be, are questions to which we shall now try to find an answer. He begins his proof by listing some propositions, which I paraphrase (p.71):

A: S is a 3020-fold sample from a raven population.

B: At least nine-tenths of the samples match the parent population to within 3%.

C: S is a sample that matches to within 3%.

D: 95% of the ravens in S are black.

E: The proportion of black ravens in the parent population lies between 92% and 98%.

What he wishes to prove is that $P(E/A.D) \geq 0.9$, since this would be an inference from a sample to the population. In statement (85), which I paraphrase, he claims that the probability that S matches within 3%, given that S is a sample and at least nine-tenths of the samples of 3000 or more match within 3%, is ≥ 0.9. (p. 72) In symbols this is the claim that $P(C/A.B) \geq 0.9$. He then shows, in statement (94), that if $P(C/A.B) \geq 0.9$, then $P(C/A) \geq 0.9$. We then have statement (95): $P(C/A.D) \geq P(C/A)$; that is: $P(C/A.D) \geq 0.9$. (p. 73) This reads: $P($ S is a sample matching within 3%, given S is a 3020-fold sample and 95% of the ravens in S are black) is ≥ 0.9.

This is a crucial step in Stove's proof, and very clearly it is invalid. If we take the direction of matching to be from the population to the sample S, then the foregoing probability statement makes no sense. If there is a known proportion of black ravens in the population and 95% in the sample S, then either S matches to within 3% or it does not. So it makes no sense to speak of a probability here. Clearly, what is meant is that there is a probability ≥ 0.9 not that the sample matches the population but that the population matches the sample. But the sudden switch is justified not at all by the previous steps, the numbered statements all being concerned with a sample matching a population, not the other way around. Once this change has been smuggled in, the rest of the proof will follow as a matter of course until we arrive at statement (102): $P(E/A.D) \geq 0.9$, this being the concluding statement of the proof.

Rejection of probability solutions to Hume's problem

Let us now submit an interim report. The first point is that if Stove's proof appears to have substance, at least initially, that is because it relies on an unexamined use of formal

probability theory, an omission which enables him to equivocate between logical and empirical meaning in his use of signs. Upon examination, however, it becomes apparent that a proof of the legitimacy of induction must operate over terms with a fixed empirical sense, which they convey along a chain of reasoning; otherwise, the argument ends with symbols devoid of content or with a change of meaning concealed by formal manipulation of terms, as in the case of Stove's attempt at a proof. Indeed, a constant theme in this first part of the thesis has been that of the confusing and illegitimate conflation of the logical and the factual, this being facilitated by the dual use of signs, as with the language of geometry, and also exhibited in the present case. Here we have had to rescue a huddle of bewildered ravens trapped, as it were, inside a spurious proof of colour correspondence between sample and population. All told, then, we are now in a position to pronounce judgement on logical proofs, which do not work, and more generally to conclude that the methods of logical probability do not lend themselves to a solution to Hume's problem.

All-pervasiveness of inductive inference

Despite this foreclosing of such possibilities with regard to a solution, one's certainties about the world remain intact, as does the need to justify them, or in the present case to justify the inference from sample to population. If I am blindfolded and allowed to select at random a hundred ravens from a flock of two hundred, and if, sight restored, I see that all in the sample are black, then it would be reasonable to infer that this is the colour of some of the others, even if I believe that Stove's proof to that effect is invalid. In fact, the systematic approach we took to the drawing of balls from a bag would also work for ravens, provided their wings were clipped, so that at least at one level the inference to their colour could be justified. This is not, of course, a level at which inductive scepticism could be countered, for if the sceptic is correct then probability theory has no legitimate application, given that any talk of its quasi-deductive character is to be dismissed, so that all that we are left with is the appeal to induction..

What I shall now prepare the ground for showing, as a crucial step in the direction of a solution, is that sceptical arguments about induction extend much further than is realised, even by their proponents. This may seem to consolidate them, but I shall argue that the gain in territory takes one in a circle, as it were, a consequence of which is that the sceptic is forced into a rearguard action against their implications. Let us begin with the connection between probability and frequency, by which I mean frequency in a series of trials, the role of which is at least partly that of confirmation. Thus, if I estimate that the probability of heads

with a particular coin is half, and if it then transpires that the coin is strongly biased in favour of tails, then I revise my estimate to bring it into line with frequency. Therefore, my original statement implied a frequency of about 50% heads, this being part of my understanding of probability.

Could it not be said that the role of frequency is just to provide evidence or confirmation of probability, itself to be regarded as being essentially different in character? This, of course, is precisely what the *a priori* theorist would say, but I have argued that it is not correct. Milk on the doorstep is evidence of a milkman, the only connection being that the one delivered the other; but the probability connection with frequency seems to me to be far more intimate, not only causally but conceptually. The reason, I think, is that rational choice is intrinsic to probability and is grounded in proportion and frequency when placed in a frame of reference.

Having just mentioned them, let us consider the connection between proportion and frequency, for instance with regard to 75 black balls and 25 white in a bag, the corresponding frequency being about 75% black balls in a series of trials. Could it be argued that the connection is evidential, the proportion of black balls in the bag being inductive evidence of a corresponding frequency? The idea here is that there is a large body of statistical evidence about this connection with regard to balls in bags, coins, cards, dice and so on, all of it supported by physical evidence, for instance with regard to the centre of gravity of a standard coin. Clearly, it is indeed the case that the proportion of colours is inductive evidence of their frequency, the importance of which is revealed when the evidential link is absent, as with a previous example of a ball in a bag being marked with a number between 1 and 100. I argued at the time that it is only within a frequency context that it makes sense to say that the probability of not-1 is greater than that of 1. There is, however, a particular difficulty here, and it arises, in the case of frequency, from epistemic reliance on numerical inequality, as further examination of the present example will show.

Suppose, then, that I bet on a particular colour being drawn. Because this is a standard probability case, it is clear that black is the rational choice, and the reason is that there is a higher proportion of black among the equipossible alternatives. If we now ask how this reason is itself to be understood, the initial answer, this being a standard case, is that we should choose the black because the equipossible alternatives are also equiprobable, so that the probability of black is greater than that of white. This brings frequency into play, in consideration of which we shall take it for granted that if a trial consists in a ball being drawn with replacement, then in the majority of sufficiently extended series of trials the frequency

of black balls drawn would be about 75%. This, however, is where the difficulty lies, for how exactly does this percentage legitimise my choice of colour? My choice is rational, or so I have claimed, because in a series of a hundred trials, say, I could expect about seventy-five black but only twenty-five white. Could it not be said, however, that I am in the same state of nescience with regard to a hundred trials as in the case of a hundred balls in a bag? The colour proportion is the same and the reliance on epistemic neutrality is the same, for I have nothing to choose between the balls in the bag, nor between the trials. And yet, we have claimed that it is the link with frequency that turns the equipossible into the equiprobable, so that probability and rational choice can be brought into play; hence our rejection of the view, which perhaps is commonly held, that purely on the basis of equipossibility it would be rational to expect black as opposed to white, or not-1 as opposed to 1 in the case of the numbered ball.

What exactly, then, is the significance of frequency in this connection? One answer is that it represents the difference between mere possibilities, as with the numbered ball, and sequences of real events, as with a series of trials, in which proportions are actualised. No doubt this is correct, but the problem is that it fails to enlighten, and also that it conceals certain facts. One of these is that all that happens, at one level, is that I take a ball out of a bag, look at it and replace it, shake the bag and then repeat the steps a certain number of times. But this is a description of uninterpreted actions, not an account of the observation and reasoning on which the expectation of a particular colour is based. What that account requires is that I engage with the notion of numerical proportion if a probability measure attaches to the expectation, a requirement which also extends, though perhaps more loosely, to what informal inductive inference leads me to expect, as when I say that I expect a black ball because more blacks than whites have been drawn. Either way, a difficulty remains, one that is reminiscent of those exploited by the inductive sceptic, who queries, after all, whether the prediction of a particular event can be grounded in one's experiences of it, the accumulation of which is normally taken to secure the inference from past to future. Thus it is that probability and induction are intimately intertwined, but always with Hume's problem looming in the background.

That said, I think that it may profit us if we ignore the problem for the moment and approach the question of frequency from, as it were, a position of sceptical innocence, so that our concern is not with Hume's arguments but with an enquiry into the role of frequency with regard to probability and rational choice. To narrow the focus in this way is not to remove the aforementioned difficulty, for we may still be puzzled as to how it is that numerical factors

can determine our predictions, the kind of question, perhaps, that might have struck the young Hume and set him on the road to scepticism. Setting aside such speculation, let us now seek answers from the discussion of an imaginary scenario, one that has more dramatic potential than the drawing of balls from a bag.

Suppose I wish to cross a border at a place where adjoining gardens of the same size have been mined, the sappers having finished in garden A but not in garden B, where only one mine has been laid. If there are a hundred mines in garden A, then the density of mines is a hundred times greater than in B, hence the rationality of my deciding to tiptoe across garden B. The point to be emphasised is that this is a rational choice in the everyday sense in which we decide on one course of action rather than another; but at the same time it still seems that our choice has to be validated in terms of numerical inequality. This suggests that the analysis is too narrowly focussed on ratio and needs to widen its scope. After all, we can alter the example so that garden B is free of mines, freeing us at the same time from any doubt as to how our choice is to be justified. How can the introduction of one little mine turn a solid decision into a choice depending on fractions? If it can, then it seems that this reasoning can be stood on its head by our decreeing that a garden free of mines is a garden with zero mines, so that our decision is based on the inequality between zero mines and a hundred mines. This is true in one sense but we need to be clear as to what it is, for the problematic sense is that in which the inequality is purely numerical, so that it holds between zero and a hundred, a difference which says nothing at all about rational choice with regard to venturing into a minefield. There is a question here, and to answer it we should return to the notion of a system.

First of all, there is nothing we need resist in the claim that we should make our way across garden B because it is easier to avoid one mine than it is to avoid a hundred of them. When it is put like that, what strikes us is that we choose garden B because it is safer, a reason which appeals implicitly to the whole of our experience of unequal sets of dangerous objects. A quiet road is for that reason safer than a busy road, the same being true of a garden with fewer mines. To cite such reasons is to place frequency considerations in the real world where our knowledge of minefields is integrated into the unified system of our general beliefs. We choose garden B because garden A is more dangerous, where 'dangerous' is a dispositional term having the implication that a risk is incurred not just on the one occasion but on all such occasions. When we speak of risk we mean that we would expect more casualties from garden A than from garden B, if the mines were replaced every time someone

stepped on one of them. Thus, our statement is verified in terms of frequency but also we may invoke frequency as evidence in support of it

To appreciate the wider significance of these facts, let us return to the connection between rational choice and frequency in cases involving balls in bags, making use of the insights we have gained from the example of the minefield. If we suppose that a bag contains 75 black balls and 25 white, so that the probability of black is 75%, corresponding to the frequency of black balls drawn, then this is to assume that the balls retain their colour when placed in the bag, an assumption which is clearly inductive in character. But now, the same applies to the whole set of assumptions which frame this probability ascription, including, for instance, those by virtue of which the alternatives are taken to be equiprobable, not to mention its being taken for granted, when drawing the balls, that one's memory of what is happening from one moment to the next is correct.

What this example suggests is that if I make a probability judgement of the kind in question, then I do so within an epistemic system which already engages with the concept of frequency having evidential value, that concept being presupposed in my understanding of the initial conditions. This has a bearing on the apparent discontinuity between present evidence and remote conclusion. In the case of 75 black balls and 25 white the present evidence is the fact that the balls have the given colour proportions, and the seemingly remote conclusion is that it is reasonable to expect to draw a black ball. This conclusion is, however, intimately connected with the evidence, for if the present colour of the balls is evidentially significant, then it is expected that their colour proportions will continue to hold. It is within a framework of such assumptions that probability is ascribed on the basis of observational evidence.

Let us now turn to the question of rational choice in non-standard cases in which frequency is unknown. If the ball in a bag is red, white or blue and I can choose to bet on blue or not-blue, what evidence can I adduce in support of my claim that probably the ball is not-blue? My only evidence is that blue is one colour and not-blue, in this case, disjoins two colours; but this evidence is not re-usable and the trial in which it plays a part is not repeatable. Or, rather, it is repeatable only if at each trial I am free to choose at random a colour complement instead of being restricted to not-blue – but this is to switch from the probability of a colour complement to that of guessing it correctly, in which case we have a standard probability experiment.

It is clear that probability, inductive inference and frequency are interconnected within a system even at the most basic epistemic level: that at which persisting physical objects are

re-identifiable and give coherence to our awareness of the world from one moment to the next. Thus it is that to rank minefields in order of safety depends on the grasp of certain concepts, for instance that of a mine, which itself is possible because mines have distinguishing features and dispositions which are frequency correlated, as are the properties by which a mine is recognisable as a physical entity, at least in its unexploded state.

It may be recalled that Hume's problem has been set aside for the moment, our present concern being to resolve a puzzle about frequency, proportion and rational belief. If we seem to have succeeded in this, it is only because we have taken a common-sense view of the matter, whereas Hume's genius was to penetrate beyond the mists of common sense to the far shore of scepticism. By way of rejoinder to our talk of a system, he would argue that epistemic interconnectedness makes inductive inference even more impossible to justify – or would do if impossibility admitted of degrees. If, for instance, it may be queried whether my car having started on previous occasions warrants the insouciance with which I now turn the ignition key, then surely the question becomes even more pertinent if it can be shown that my inductive premiss depends on assumptions which are themselves inductive in character. These concern keys and cars being re-identifiable and stable in appearance and function; also, it is obvious that I take for granted, with regard to past occasions on which the car has started, that my memory of these successes is veridical. But if the Humean sceptic challenges me on this, asking how I know that my memory is correct, then one response, surely, is to appeal to the general reliability of my memory, itself an inductive argument, one which I may now be asked to justify in its turn, along with my claim as to having remembered correctly in the past. That, or so I surmise, would be Hume's rejoinder – but I also think that he would be driven to it, for I imagine that he would anticipate the mists parting to reveal an armada of anti-sceptical arguments bearing down on him.

Inductive scepticism entails global scepticism

The source of this threat just is the fact, as I shall now try to show, that consistency demands that inductive scepticism be deployed against observation statements and the deliverances of memory, contrary to Hume's insistence on restricting its scope. The picture that Hume presents – one endorsed by Stove – is of a world exactly like our own, containing birds and balls and bags presently observed or belonging to the past, all beyond the reach of the sceptic except in one respect: that we have no reason to believe that the present constituents of this world will continue from this moment on. My aim is to show that this is not a coherent

picture, and the place to start is with the sceptic allowing safe passage to our knowledge of the past.

Let us first of all take unobserved events to include future events as well as present events not observed, so that an observed event is taken to include observed past events as well as events now being observed. Then we are querying whether one's beliefs about unobserved events are problematic in a way that the corresponding beliefs about observed events, in this case past events, are not. If, to begin with, the sceptic argues that predictions about the physical world are non-demonstrative, then surely the same is true of direct memory or of the reasoning, whatever the form it takes, by which one establishes that a past event has occurred. Secondly, it is also clear, as the previous paragraph attests, that in a particular case of justifying memory we may appeal inductively to its general reliability. Thirdly, it is taken to be a solid ground for trusting in my memory that my memories are strongly confirmed when predictions or expectations prove to be correct. For instance, I seem to remember that I placed nine black balls and a white ball in a bag: if, for whatever reason, I wish to confirm the veracity of this memory, and if I do so by looking in the bag, then this, courtesy of the no-miracles argument, is very strong confirmation, for how else is it to be explained that I was correct in what I expected to find? In different circumstances there would, of course, be other possibilities, for instance that it was by way of testimony, not direct inspection, that I knew what the colours would be. But then my looking in the bag would confirm my memory of having been told what the colours were, and again this would be very strong confirmation. So strong, indeed, that I shall argue in a later chapter not only that our knowledge of the past is grounded in confirmation of this kind, but also that it is essential to the very concept of the past. For our present purposes it is enough to point out that if we have no reason to believe in the future, then we have a much reduced reason, in fact no reason at all, to believe in the past. As a final nail in the coffin, we note first that Hume rejects the induction principle or that of the uniformity of nature on grounds of circularity, thereby demolishing a counter-argument against his sceptical position. Similarly, if I make *any* attempt to counter the sceptic about knowledge of the past, then I am guilty of begging the question, for reliance on memory is essential to the declarative use of language.[5]

If this is correct, so that the inductive sceptic *must* extend his reasoning to our knowledge of the past, then we need to ask whether it also applies to observation reports, the answer to which will depend on what is meant. If they are taken to be descriptions of the

---

[5] See chapter 6.

physical world, as in the probability example just given, then they have inductive implications which make them vulnerable to the sceptic. When I say that this is a ball drawn from a bag, or just that this is a ball, it is not as if the truth of my statement about a present object is compatible with any future state of affairs whatsoever. I imply, for instance, that the object is three-dimensional, this being a condition of my correct use of the word 'ball', and part of what is implied is that the object will look more or less the same, being spherical, if I turn it in my hand or walk around it. If I now describe the object by its colour, saying that it is black, then I imply that this is a stable property; also, that other people will agree with my description if I ask them to tell me what they see. It is true that these are simple conditional statements rather than statements in the future tense, but this does not invalidate the point, for propositions of this kind are predictions at one remove. Since they take the form, 'If A occurs, B will occur', they imply the possibility of predicting B on the basis of A. Also, there is a degree of interchangeability here: if simple conditional statements go unchallenged by the sceptic but statements in the future tense are turned away, they can be smuggled through by having a simple conditional clause added to them. Clearly, if I have no reason to believe any statement about the future, then I have no reason to believe any simple conditional statement. So it would seem that inductive scepticism is not consistent with physical description: if I cannot say what this ball would look like, then I cannot say what it is.

Given Hume's uncritical use of physical object language, for instance when referring to the game of billiards in his analysis of causal concepts, it would seem that he is guilty of inconsistency, a charge considered by Ayer, as we shall see in a moment when we turn to what he has to say. A more lenient view is that Hume is forced into acceptance of his irrational belief in the existence of things, or that he takes it that succumbing to such irrationality does not detract from the cogency of his sceptical thesis: that inductive inference is not demonstrative and can be otherwise justified only by arguing in a circle. To ask whether there is a sense in which he would be correct in this, so that there is no such detraction, let us return to the stock example of the rising of the sun. If I say that the sun has risen every day in the past, therefore it will rise tomorrow, then it is clear that the premiss may be asserted and the conclusion denied, so that the relation between them is contingent in this particular case. This helps to consolidate the relevant part of Hume's thesis, despite the fact that to tease out its finer points we should have to explicate our own use of 'demonstrative' in order to keep faith with exactly what he meant. This leaves the other part, about arguing in a circle, which applies to a more realistic view of what we should say in the present case: not categorically that the sun will rise but that it probably will. As against this,

Hume would maintain that the appeal to probability presupposes induction by assuming that the future will resemble the past, so that inductive inference cannot be based on it.

If we now turn to the sense in which Hume may be correct about the cogency of his arguments not being affected, the possibility here is that his sceptical position is able to establish itself in terms which in another context those arguments would undermine. In the present case this is to say that the arguments show that if the present and past existence of the sun is assumed, then this gives no reason to believe that it will exist in the future, a fact which continues to hold even if, by an extension of those arguments, there is no reason to believe that the sun exists, given that to speak of the sun now shining is to imply constraints on future possibilities. The correct rejoinder here, I think, is not to object to this reasoning but to point out that the sceptic can be consistent, without which he has no case, only if he follows his arguments through, even if it is to the conclusion that we have no reason to accept observation reports as these are ordinarily understood. If his arguments are taken to this extreme, and if anything is to remain of his factual beliefs, however attenuated, then it is clear that observation statements will have to be construed in such a way as to present less of a target. Since this can be achieved, if at all, only if our descriptions are confined to what is given, we now have to enter the world of sense-qualia and immediate experience.

Intrinsic description

To that end, we should perhaps allow Ayer to be our guide, given that he is regarded as one of the inheritors of Hume's empiricism, though in a form which is very much nuanced by his own point of view. Speaking of the thesis that factual inference is not demonstrative, which Hume presents in terms of cause and effect being distinct events, Ayer tries to show that the thesis has a wider reach than the causal one, and he does this by introducing, as follows, the notion of an intrinsic description:

I shall say that a description of the state of a subject S at a particular time t is intrinsic to S at t if and only if nothing follows from it with regard to the state of S at any time other than t, or with regard to the existence of any subject S^1 which is distinct from S, in the sense that S and S^1 have no common part. (1972, p.6)

Following Ayer, let us call the object of such a description a distinct event. We have seen that even if events are not distinct in this way, they may still be only contingently connected, as with the connection between different sunrises, or indeed between any events that are suitably described. The point of the definition must therefore lie elsewhere, and it seems that Ayer's concern is with what he refers to as Hume's atomism. It is true by definition, he says, that 'if

two events are distinct in this sense, an intrinsic description of either one of them entails nothing at all about the existence or character of the other' (Ibid). Hume's thesis is now taken to be the ontological one that there are indeed events that are distinct in this sense, such that, to quote from Ayer, 'they are sufficient to describe everything that happens' (Ibid). He thinks, too, that intrinsic descriptions have to be purely phenomenal, applying, for instance, to colours and shapes, though only if less is meant than is normally the case. Using Ayer's own example, when I say in an everyday context that this table is brown I imply that it is disposed to look brown in certain conditions, such that it need not at present look brown to me. Ayer thinks, however, that in theory these dispositional properties could be replaced by occurrent ones, a description of which would be intrinsic in the required sense. He believes, too, that this does not commit one to phenomenalism, for it is,

possible and legitimate to start with neutral sense-qualia, and then represent statements about physical objects as elements in a secondary system which functions as a theory with respect to the primary system of sense-qualia. Then, Hume's point can be put by saying that statements in the secondary system are not deducible from statements in the primary system and that these in their turn are not deducible from one another. (p.9)

If he is correct, then these latter statements fit the definition of intrinsic descriptions, which therefore gain purchase at the level of sense-qualia.

Ayer is hesitant about some of what he says, and he wonders, for instance, whether the possibility of starting with sense-qualia as primary elements depends on our being able to locate them in time and space. The thrust of his empiricist thesis is definite enough, however, and it conflicts with our own views on the workings of a system. If Ayer is correct, then although Hume may be inconsistent at one level, his basic sceptical position still holds, or would do if he retrenched at the deeper level of sense-qualia, the level at which re-assigned descriptive terms embrace momentary particulars. If this were achievable there would be primitive information beyond the reach of the sceptic, though it would bring little comfort to those who cherish their everyday beliefs, such as that their bodies exist, or that when they utter a sentence they know what they have said. The sceptic, on the other hand, would still have much to be pleased about, for his anti-inductivist arguments could prevail over factual beliefs whatever one's philosophical view of them. Thus, the phenomenalist has no epistemological advantage over, for instance, the naïve realist. If factual statements translate into ones about actual perceptual sensations and those to which we infer, then clearly the inference is vulnerable to sceptical assault; and this would also be true on the naïve realist view, where the inference is from the present to the future existence of physical objects. Also, if factual statements are theoretical with respect to primary evidence in the form of sense-

qualia, as Ayer suggests, then again it is obvious that Hume's arguments could be deployed against the inferential aspects of the theory.

The question we now have to ask, however, is whether it is really true, with regard to statements that refer to distinct events, that 'they are sufficient to describe everything that happens', and whether they are proof against the arguments of the sceptic. I think that we have to answer in the negative in both cases, and the reason, as I shall now try to show, is that the very notion of an intrinsic description is incoherent, because factual knowledge forms a system, this being a more appropriate metaphor than that of an epistemic structure erected on sensory foundations. To argue in this way and to trace out its implications is, however, a large undertaking, at least in the context of a wider discussion of what it is to mean and to understand and of conditions for the use of declarative sentences, all of which is relevant in various ways. That discussion will extend into the next two chapters, the subject of which will be Wittgenstein's theory of meaning and understanding. What I propose in the meantime is to prefigure the arguments by showing that Ayer's definition has no application, this to be achieved by further examination of his views, sometimes in depth but also by skating over contentious issues, prior to engaging with them at a deeper level in later chapters.

Let us begin with Ayer's acknowledgement that the subject of an intrinsic description 'can only be private, fleeting sense-impressions, which will presumably have to be identified by demonstratives. So our primitive statements would have the form "this is f" where "f" is an intrinsic description of the sense-datum designated by "this" ' (p.9). An example would be 'this is blue', which grammatically is a declarative sentence; but it is not thereby guaranteed to pick out an object and attribute a property to it, for 'hist is lube' is also a sentence, one that means nothing. What is also needed is that the use of the sentence be rule-governed, a question immediately arising as to whether this condition allows the sentence to convey an intrinsically descriptive statement. A basic rule is that to endow 'this is blue' with meaning I must use it to refer to only what is blue. This is obvious if it is imagined that I am promiscuous in my use of 'blue' and attach it to any colour, not just that one. It follows that 'this is blue' has semantic content only if I am disposed to use it in a particular way, in this case to refer to my sense-impressions of blue. An interesting question now arises as to my cognisance of this condition, and it is arguable that I must *think* that my use of 'blue' is meaningful, for it would make no sense for me to say that this is blue but that I am not disposed to use 'blue' correctly. On the contrary, if I say that I know what I mean by 'blue', or what is meant, then I imply, for instance, that if an assortment of colours includes blue, then I shall be able to identify it as that colour. But this is a projection into the immediate

future – it has to be immediate, for in order to understand or give meaning to 'this is blue' on any particular occasion I am not required to be suitably disposed on a permanent basis, or even for an hour, depending on what happens in the meantime. It is possible, after all, that during that period I lose my grasp of what the colour word means.

None of this is to deny the obvious sense in which blueness predicated of sense-impressions may be occurrent, so that evidence and conclusion coincide, the evidence that this is blue just being my experience of it as blue; hence the claim that descriptions of this kind are incorrigible. My point is that even if such colour events are occurrent, it still is a pre-condition of the correctness of the appropriate predicate that one should be suitably disposed with regard to its use, so that in this sense one's description always goes beyond what is given. Ayer's mistake, perhaps, is to take the relevant dispositional and occurrent properties to attach to objects, not to our understanding of them. Thus, he speaks of colour predicates going beyond what is given, so that they are not intrinsic in their ordinary use: 'For this table to be brown, it is not sufficient and indeed not necessary that it now looks brown to me, or to any other given observer; it has to be disposed to look brown to most observers under such and such conditions, perhaps also to reflect light of such and such a wave-length.' (p.8) What this overlooks is that nothing can be a colour predicate unless the use of it satisfies criteria of meaning and understanding, without which it is an empty sign, as 'brown' would be, with no connection to such facts as that a table is disposed to vary its appearance according to how it is lit. These criteria are partly dispositional, so that my understanding of 'brown' depends on my being suitably disposed in my use of it. The point is just that this condition continues to obtain even when my use of the word relates only to subjective colour impressions at a particular instant. If this is correct, then the claim to incorrigibility should be rejected, or re-interpreted[6], as should the view that intrinsic descriptions gain a foothold at the level of what is immediately given in perception.

Let us press these arguments home, concentrating this time on the fact that by definition an intrinsic description applies only to the state of a subject at an instant, not to other subjects or instants. The first point is that Ayer equivocates in his use of physical world terminology: he stipulates that intrinsic descriptive terms should be silent about the state of the subject at other times, thereby implying that it may be re-identified, a possibility which the definition banishes by decree. He appeals to the use of demonstratives, as in 'this is blue', but again he has to reach up into the physical world, where the conventional use of 'this' as a

---

[6] See discussion of avowals, chapter 5.

grammatical subject is to label re-identifiable objects under particular descriptions. It is true that there are grammatical rules which allow continuity of reference, so that an object described as it now is can be tracked by language as it slips into the past. Thus, 'this is blue', if its object is a short-lived event, becomes 'that was blue', both sentences referring to the same instance of blue. Clearly, however, an intrinsic description is not able to avail itself of this facility, or not if the sceptic is to be kept at bay, for 'that was blue' refers to a past event, the evidence for which is non-existent if the sceptic is both consistent and correct. But now, if an event described intrinsically is instantly swept away and lost, so that the description lapses, then we have to ask whether there is any sense at all in which the description captures the event. One could go on in this vein, for there is much to be said along these lines, including, for instance, its being pointed out that Ayer's definition incorporates the concept of time, again raising a suspicion that physical world concepts are smuggled in at the same time as their sense is apparently stripped away or pared down.

One way of altering the angle of attack, as it were, is to view all description, whether intrinsic or not, as having physical and mental shape, for example when 'this is blue' is uttered or thought. In the following chapters I shall argue that one cannot grasp a statement unless one is aware of it as a semantically charged sentence, such that one's recall of it as part of the process by which it is read, spoken, written or thought is not in question at that time, or is queried only if in the same way one's recall of any expression of doubt is itself taken for granted. If this is correct, so that reliance on memory is implicit in one's use of words, then again it follows that observation or experiential statements, whatever their analysis, go beyond what is given.

Perhaps it is time to recapitulate the main points of this chapter and to look beyond them to the next. A notion which our rebuttal of probability solutions has thrown into relief is that of probability applications forming a system, its component parts feeding into one another, such that the link between frequency and rational choice is implicit in any statement of initial conditions, for instance that a bag contains such and such a proportion of red and blue balls. But frequency and induction are cogs within a system in which they engage with language, the whole of it lubricated by memory[7]. A corollary of this interconnectedness is that inductive scepticism is forced into a rejection of all factual inference, hence of empirical knowledge, and not even the simplest experiential report gives respite to the sceptic in his

---

[7] There are non-linguistic epistemic systems, as with bats, but our concern is with systems that allow for scepticism, essential to which is the use of language found only in our own species, unless other creatures are afraid to say anything.

descent into solipsism. But now, if Humean scepticism has to be transformed in this way, and if we are correct in our outline account of the way in which induction, memory and disposition enter into meaning and understanding, then a vista opens up which contains the possibility of proving that Hume's anti-inductive arguments are self-refuting. It all depends on whether the position we have outlined can withstand scrutiny, and on whether we can arrive at a satisfactory theory of what it is to mean and to understand. These are the questions I shall now try to answer.

Chapter 4

Wittgenstein On Meaning and Understanding

What I propose in this and the following chapters is to consolidate certain arguments which so far have been only sketchily presented and to work them up into an anti-sceptical theory derived from the notion of a system. The view which is now taking shape is structured around the thesis that if inductive scepticism threatens knowledge of the past and of the physical present, then this may be combined with certain facts about the interconnectedness of empirical inference and linguistic use, thereby yielding the conclusion that the sceptic's own use of language is fatally compromised. Whether this anti-sceptical edifice can be made imposing enough will depend on the view we take of belief, meaning and understanding, the exploration of which we are about to embark upon. In the previous chapter the main point made in this connection, in line with common sense, was that there are dispositional criteria of understanding and of other forms of intentionality; however, common sense would also have us believe that at least some of these forms, such as understanding, are occurrent in their manifestations, a question now arising as to how the two sides to their nature are to be married up. One may ask, too, whether understanding is a mental state, and what it is to grasp a rule, linguistic rules being of special interest, as are arithmetical rules in the case of Wittgenstein's explorations in this field.

My procedure, as before, will be to work towards a refutation of scepticism by critically examining other people's views, in this chapter mainly those of Wittgenstein, whose *Philosophical Investigations* will be our primary source. Since much of what he says is relevant to the problem of other minds, I shall comment on that aspect of it as we go along, and then refer to it again in the other minds chapter; but my main concern in this one will be with the problem of induction. He has little to say that is directly about induction, and makes no mention of Hume's problem as such; but he does concern himself with intentionality, which we know to be relevant to the problem. What he says more directly, when speaking of one's certainty that one is able to continue a series, or that an unsupported book will fall, is that in neither case do we need any grounds, and also that nothing could justify our certainty better than success. He speaks of justification or reasons coming to an end, and he places emphasis on the way in which the word 'justification' is used. He would claim, perhaps, that the way in which we reason forms a practice, within which the word 'justification' has a use,

and would thereby imply that it is misconceived to ask how the practice itself is to be justified. (1958. Sections 324-5, 481 and 486)[1]

None of this would satisfy Hume, who quite possibly would find it incomprehensible, along with Wittgenstein's theory of intentionality. As for grasping all this ourselves, or the philosophy in which it is embedded, I suggest that we examine in detail key passages in the *Investigations*, one advantage of which is that we can make our own exegetical mistakes instead of having to rely on other people's. This close attention to the text is to some extent imposed on us by the difficulty, not so much of summing up the author's views, which are opposed to epistemological scepticism and to the very notion of a genuine philosophical problem, as of distilling his arguments from the flow of gnomic utterances by which he conveys them. Sometimes, however, the task will prove too much for us, and then we shall seek enlightenment from his commentators.

Perhaps we may begin with Wittgenstein on understanding. He supposes that there are two people, A and B, and that B watches A write down the numbers 1, 5, 11, 19, 29 as the initial terms of a sequence, at which point he says that he knows how to go on. Wittgenstein now asks what it is that B's understanding consists in, and he suggests several possible answers. Perhaps B tried the *n*th term formula $a_n = n^2 + n - 1$ after A had written the number 19, so that the next number confirmed his hypothesis. Wittgenstein points out, however, that the formula could occur to B without his understanding it:

For it is perfectly imaginable that the formula should occur to him and that he should nevertheless not understand. "He understands" must have more in it than: the formula occurs to him. And equally, more than any of those more or less characteristic *accompaniments* or manifestations of understanding. (152)[2]

And again:

If there has to be anything 'behind the utterance of the formula' it is *particular circumstances* which justify me in saying I can go on – when the formula occurs to me. (154)

Let us consider the thrust of Wittgenstein's argument. He has yet to specify what the 'particular circumstances' might be, though he promises to shed light on them when he investigates the concept of reading, this being the task that he sets himself next. In the meantime, the following comments may help to fill the gap. Suppose that several *n*th term formulas, including $a_n = n^2 + n - 1$, occur to B as flotsam in his stream of consciousness, so

---

[1] The question of knowledge and justification is explored more fully in Wittgenstein's *On Certainty*, to be discussed in chapter 6.
[2] In quoting from *The Investigations*, I shall refer only to section, not page, numbers. Also, my use of single and double quotation marks will be the same as in the quoted passages.

that he barely glances at them as they float past. Then if he is now told the formula, or otherwise comes to know it, and if he remembers it passing through his mind earlier, he does not claim that at that earlier time the understanding of the sequence came to him, despite the fact that the formula did. There is a distinct difference between this earlier occurrence of the formula in his mind and the way in which he now recognises it as the formula for the sequence – or at least there *seems* to be. The point about this difference is that it concerns an occurrent mental event; what also has to be said, however, is that such an event could not *in itself* constitute understanding, since we need only *think* that we understand for it to occur, and also for the reason that to understand anything at all is to have the appropriate dispositions.

Reasoning as an occurrent, interior process

Clearly, then, we can agree with Wittgenstein that there are interesting questions about what it is to understand, if that is what he would say. In the same breath as he speaks of a formula coming to mind, however, he also mentions trying out different ones and using them to predict the next term. This is very dissimilar to the case in which a formula is merely glanced at, and the main difference is that in trying out formulas we engage in a process of constructing and testing hypotheses, to which end we use inference and calculation. Perhaps, then, I may venture even at this early stage to provide a corrective to a particular construal of what it is to understand. It is clear that to think of a formula or write it down is not necessarily to understand it, for it may present itself to us as devoid of arithmetical meaning, for instance as a string of unfamiliar symbols, or it may be recognised as a formula but not as *the* formula, the one applying to a particular sequence or series, or it may be mistakenly recognised as such. And, as just mentioned, it cannot in any case in itself constitute understanding. It would be wholly inaccurate, however, to speak in this way of conscious *reasoning*, as in the case of trying and testing hypotheses, if this is to imply that the process has to be overt in order to manifest understanding, or that it is not enough in itself to be an instance of what it is to understand. If, for instance, I try out various nth term formulas for a particular sequence of numbers, the process being overt only insofar as I am looking at the numbers written down, then it may be that I mentally test each formula by assigning values to its variables, these values being the numbers in the sequence, or that I examine the sequence for numerical regularities in order to work out a formula, and that it seems to me that I understand that this is what I am doing, and that I understand it *now*. An exactly parallel case is that in which I read an English text and understand it as I read it, so that my understanding

is immediate. I may be wrong, of course, about my grasp of particular words, just as I may have calculated incorrectly in the numerical case; but the point being made concerns the nature of understanding, not the question of whether one understands. Besides, I cannot later come to doubt, except in the most bizarre circumstances, of which I had no awareness at the time, that I understood most of what I had read, or most of the process of working out the formula. Mistakes are always possible, but this does not detract from the view that subjective cognitive processes are essential to understanding. We shall have to see how this impinges on Wittgenstein's thesis equating meaning with use or on his claim that inner thought processes stand in need of outward criteria, with its implications not only for meaning and understanding but also for the evidential aspect of our knowledge of other minds. In the meantime there is what seems to be the significant fact that one's understanding may be exhibited to oneself in one's reasoning, in this case in the form of working out a formula, a process the grasp of which seems immediate and occurrent, so that it is this that we should focus on rather than on seeing or imagining a formula in isolation.

Wittgenstein on reading

Let us now turn to Wittgenstein's investigation of the particular circumstances attaching to the ability to read. By 'reading' he primarily means reading aloud but without necessarily understanding the words, since one may know how to pronounce them without knowing what they mean. As far as one can tell, he is sceptical of the assumption that reading is a special conscious activity of mind, such that 'the one real criterion for anybody's *reading* is the conscious act of reading, the act of reading the sounds off from the letters' (159). Thus, he asks us to imagine that:

A wants to make B believe he can read Cyrillic script. He learns a Russian sentence by heart and says it while looking at the printed words as if he were reading them. Here we shall certainly say that A knows he is not reading, and has a sense of just this while pretending to read. For there are of course many more or less characteristic sensations in reading a printed sentence; it is not difficult to call such sensations to mind: think of sensations of hesitating, of looking closer, of misreading, of words following on one another more or less smoothly, and so on. And equally there are characteristic sensations in reciting something one has learnt by heart. In our example A will have none of the sensations that are characteristic of reading, and will perhaps have a set of sensations characteristic of cheating. (Ibid)

Perhaps the first step in glossing this passage is for us to query Wittgenstein's use of the word 'sensation', which does not seem entirely apposite in the context of its being said that there can be a sensation of hesitating, looking closer or misreading, or of saying something by heart, rather than that one is aware of them or that they are conscious acts or states. It may be,

then, that his use of this word betrays a particular bias, a possibility to which we shall presently return. It seems odd, too, to suggest that A knows that he is reading on the basis of what is characteristic of that activity. If, in my own case, it is on that basis, then part of my evidence that I am reading is that I hesitate, correct myself, lose my place, look at the words at the same time as I say them, and so on. But this makes it sound as if on the basis of the evidence I think it probable that I am reading. The fact is, surely, that even when I say aloud only a single, short sentence, so that the hypothesis that I am reading would be only weakly confirmed, it never occurs to me to doubt that I *am* reading the words, as indeed is the case when I say the first of them. Clearly, this talk of evidence and confirmation is very ill-suited from a first-person point of view to the concept of reading, and it would be better to say not that it seems to me that I am reading but that I already know that I can read, this being the activity in which I am now engaged.[3]

No doubt there is more to be said, but our present concern is with the analysis of a concept, in this case not even of reading in the usual sense but of reading aloud without understanding. In reading aloud one is not required to exhibit any great skill, in the sense of correctly matching the spoken to the written word, and it is easy to imagine a foreign student of English reading aloud in a halting manner in which every word is wrongly pronounced. Indeed, the essence of what is involved may be illustrated in many other ways, one of which finds us yet again presented with balls in a bag, this time with a view to naming the colour of each ball as it is drawn. What is essential here is that naming the colour derives from recognising it, and if the process of deriving is generalised to encompass reading, naming and direct physical description, then arguably it is fundamental to the connecting of language with the world. This, however, is what Wittgenstein now goes on to say:

But imagine the following case: We give someone who can read fluently a text that he never saw before. He reads it to us – but with the sensation of saying something he has learnt by heart (this might be the effect of some drug). Should we say in such a case that he was not really reading the passage? Should we here allow his sensations to count as the criterion for his reading or not reading? (160)

Now it seems to me that these remarks veer uneasily between first-and third-person points of view, but at least they indicate what it is that Wittgenstein is trying to show. He is not denying that the person has sensations, or reducing them behaviouristically; rather, these remarks perhaps again prefigure his thesis that inner processes stand in need of outward criteria. If, however, one examines this particular case, and from a first-person point of view, then it is impossible to imagine having the sensations of reciting by heart at the same time as

---

[3] The later discussion of avowals is relevant here. See chapter 5.

one reads aloud from an unfamiliar text. If this is not immediately apparent, then perhaps the reason is that it is obscured by Wittgenstein's use of the word 'sensation', as if introspection when reading aloud or reciting by heart reveals only a series of sensations as the conscious content of each process. The truth is rather that I derive the spoken from the written word as part of the perceptual experience of reading, so that the reference should be not to my 'sensations' of reading but to my *awareness* that I am thus engaged. By the same token, the reference should be to my awareness of reciting by heart. What Wittgenstein is really inviting us to imagine, then, is that I read the text aloud at the same time as I think that I am reciting it from memory. But how could this be possible? All I need do is close my eyes and, lo and behold, I cease to say the words! I am no longer able to see them, whereas if I knew them by heart I could continue with the charade, at least until it was noticed that I was reading with my eyes shut.

In making these criticisms, and in the earlier paragraph, I have cited what is arguably the main defining property of the concept of reading aloud: that of deriving the spoken from the written word. Wittgenstein, however, asks of a person we have taught to read, 'But why do we say that he has *derived* the spoken from the written words? Do we know anything more than that we taught him how each letter should be pronounced, and that he then read the words out loud?' (162) The obvious answer here is that we know that in other people reading involves deriving, as it does in ourselves, and that we have as much reason to attribute this process to others as we have to ascribe pain to them in suitable circumstances. The obvious is not to Wittgenstein's taste, however, for he now turns from reading aloud to transcribing, and asks us to suppose that we show our pupil how to transcribe from print to handwriting by giving him a table with printed letters in one column and corresponding cursive letters in the other, so that he transcribes a printed text by consulting the table. We are now to imagine that 'when he did this he always wrote *b* for A, *c* for B, *d* for C, and so on, and *a* for Z? – surely we should call this too a derivation by means of the table' (163). He continues:

Suppose, however, that he does not stick to a *single* method of transcribing, but alters his method according to a simple rule: if he has once written *n* for A, then he writes *o* for the next A, *p* for the next, and so on. – But where is the dividing line between this procedure and a random one?

But does this mean that the word "to derive" really has no meaning, since the meaning seems to disintegrate when we follow it up? (Ibid)

Well, no, because the meaning is clear even in these anomalous cases. And as for his rhetorical question about the procedure being a random one, no it is not: it would be pointless and cumbersome to move on to the next cursive letter every time a printed letter was

repeated, but there would be nothing random about being so systematic. It would not, by the way, be a reversible transcription process, for any sequence of printed letters would map uniquely onto a sequence of cursive letters but not vice-versa, thereby making it impossible to translate from cursive to printed. The fact remains, however, that the procedure allows for cursive letters to be uniquely derived from printed letters. Nor does it matter that in the present case different cursive letters are made to correspond to repeated instances of the same printed letter: what is important is that the whole process involves recognising the letters and being able to derive one from another. Wittgenstein, evidently, would not agree, for he now continues:

In case (162) the meaning of the word "to derive" stood out clearly. But we told ourselves that this was only a quite special case of deriving; deriving in a quite special garb, which had to be stripped from it if we wanted to see the essence of deriving. So we stripped those particular coverings off; but then deriving itself disappeared. – In order to find the real artichoke, we divested it of its leaves. For certainly (162) was a special case of deriving; what is essential to deriving, however, was not hidden beneath the surface of this case, but this 'surface' was one case out of the family of cases of deriving.
And in the same way we also use the word "'to read" for a family of cases. And in different circumstances we apply different criteria for a person's reading. (164)

What this indicates is that Wittgenstein finds it significant that there are family resemblances rather than essential similarities between different cases of deriving, which he suggests is also true of reading. The fact is, however, that the use of the verb 'to derive' is similar in similar cases. Thus, if I read aloud a line of unfamiliar text, which is to say that I derive each spoken word from a corresponding written word, then this is what I do for each word in the line and each line in the text, a fact which raises a question as to the relevance of dissimilar cases of deriving. Perhaps Wittgenstein appeals to them in order to show that the use of 'to derive' or 'to read' will depend on the particular circumstances. I suspect, however, that he would take exception to the following as an instance of the circumstances attending a particular case. Suppose that I read aloud from a text with which I am unacquainted, so that I derive the spoken from the written word. This means that as my eyes track across the page then for the most part I do not know exactly which words are coming next, so that it is only when I see them that I am able to say them, the shape of the particular words I utter being in this sense uniquely determined by my reading of them. When, on the other hand, I cannot read but only pretend, having memorised a spoken script, then I do not derive what I say from what I see.

What this example indicates is that there is indeed a conscious activity of reading aloud, essential to which is the process of deriving the spoken from the written word. Even if we agree that on a wider view there is no essence of reading, such as to transcend the

different senses of 'to read', we may still ask whether the fact that there are different kinds of reading has anything to contribute to a discussion of any particular one of them. There is, as all shoppers know, the kind of reading in which an electronic device scans the bar code on a plastic strip, unless it is illiterate, and here there is no conscious activity at all, let alone a conscious process of deriving one thing from another. Clearly, this tells us little about a person reading aloud, so perhaps we should consider a more similar form of reading. We have yet to mention vocalised Braille reading, for instance, which differs from visually derived reading aloud only in that the spoken word takes its cue from the felt shape of the inscribed word. But now we seem to have the opposite problem, for it is this very similarity which militates against any difference, such as it is, illuminating the connection between inner and outer aspects of reading aloud.

To criticise Wittgenstein in this way is to take a common-sense approach to the consciousness of other people, as also to what it is to derive the spoken from the written word; but it is this approach that Wittgenstein continues to call into question. Focussing on the phenomenal aspects of the process of reading, he now asks, 'But when we read don't we feel the word-shapes somehow causing our utterance?' (169) He now invites us to read a sentence aloud and then to say it at the same time as we look along a line of symbols in arbitrary sequence:

Can't one feel that in the first case the utterance was *connected* with seeing the signs and in the second went on side by side with the seeing without any connexion?
  But why do you say that we felt a causal connexion? Causation is surely something established by experiments, by observing a regular concomitance of events for example. So how could I say that I *felt* something which is established by experiment? (Ibid)

 The assumption here is that we, presumably we philosophers, think that the reader somehow senses or feels a causal connection in the very process of reading aloud. This is reminiscent of Wittgenstein's earlier remarks on the sensations associated with understanding, a topic to which he is about to return. His point would seem to be that awareness of reading as a causal process is not mediated by any 'feeling' of direct acquaintance with a causal relation. Such a feeling does not permeate, as it were, the experience of reading, thus enriching its distinctive character as a conscious process, but is rather a matter of what readers say when they reflect on what they do. The fact is, however, that what they say makes sense, and it may seem to them that they are reporting a conscious feature of reading aloud. This is attested to by what we are disposed to say when we speak from a script and are asked whether we are reading it as opposed to having memorised it. If the former, and if this is our first reading, then we would no doubt say that in reading aloud we are aware of deriving the spoken from the

written word, where this is to imply a causal process. Thus, the dispositional aspect of my awareness does not preclude my issuing a direct report of the causal features in my conscious experience of reading, and the reason is that my not being able to 'feel' a causal connection does not prevent me from being aware of it at the time. This, as we shall see, is reminiscent of McGinn's attack on what he takes to be Kripke's assumption that semantic facts have to be reducible to facts specified non-semantically.

Our present concern is with Wittgenstein, whose purpose in denying that one 'feels' a causal connection is perhaps to draw attention to ways in which we as philosophers are misled. This is consistent with his rejection of reading as a special act of consciousness, the emphasis being placed, rather, on behavioural criteria of reading, these being more pronounced in the case of reading aloud than in that of reading silently, with or without understanding. None of this amounts to behaviourism in any clear sense, for Wittgenstein refers non-reductively, as far as one can tell, to feelings, sensations, mental images and so on, except that he would require these to have outward criteria. One view to have gained currency is that for Wittgenstein it is *a priori* that certain forms of behaviour are evidence of inner processes, but without logically guaranteeing them and therefore without being identical with them. Clearly, this is relevant not just to the analysis of the concept of reading but also to the other minds problem. The difficulty here is not just to tease out what could be meant but to avoid trivialising it. There is, after all, a very obvious sense in which it is only on the basis of others' behaviour and actions that I could have any conception of them as fellow human beings, ones who are, for instance, capable of feeling pain, or of reading; but this is part of what is given rather than the solution to any puzzles arising from what we know. Besides, there are significant conceptual differences between pain and reading aloud, the former being different from behaviour whereas the latter includes it. Perhaps, then, we should now consider non-cognitive silent reading.

Suppose that a monolingual foreign student is looking at an English text we have given him and that we speak only English, so that there is no verbal communication between us. Then our evidence that he is reading to himself, that he is mentally pronouncing the words, albeit without understanding, would have to be indirect, his looking at the text being only a necessary condition of reading. We could combine it, I suppose, with already knowing that he is able to read English aloud, from which we infer that he is now reading to himself. According to Wittgenstein, however, when a person reads aloud all that we know is that he is pronouncing each word of a text, rather than that he is deriving the spoken from the written word. If that is correct, then the student being able to read aloud provides no evidence at all

that he is now reading silently to himself, an activity with no unmistakeable outward sign. This, however, is nonsense, for we all know what it is to read silently to ourselves, an activity in which I am now engaged as I write these English words, thinking that I understand them, and in which I was once engaged, initially without understanding, when learning French. We believe that the foreign student is reading to himself because we know that he can read aloud, the one activity being strongly correlated with the other, and that is all there is to it. When it comes to his reading aloud, we may check that he is by following the text as he reads from it, but this itself is an inner mental process on our part, one in which we associate the written with the spoken word. Again it has no outward sign, apart from our looking over his shoulder as he speaks, an action to which a variety of interpretations may be attached.

If this reasoning is along the right lines, then the concept of non-cognitive reading, whether silent or spoken, is such as to depend on inner processes as truth conditions. If, the student now having started to utter the words, I look over his shoulder and claim that he is reading aloud, and if I do not thereby ascribe to him the inner process of deriving the spoken from the written word, then all that happens is that I associate the words he speaks with those I read, a correspondence which on its own makes no sense of its being said that he is reading, or that I am checking that he is. What, in any case, would the distinction between reading aloud and reciting by heart amount to if it picked out only behavioural differences? In this connection it is true that Wittgenstein is not a strict behaviourist, for his view is that inner processes require outward criteria, as with his remark that a formula coming to mind betokens understanding only if particular circumstances obtain, presumably those in which it is publicly exhibited. But he also says that when a person reads aloud we know nothing of any inner process of deriving, rather than that such a process must be manifested in behaviour. In this matter our wish to accommodate Wittgenstein may conflict with everyday intuition, all the more so if it is applied to silent reading, the attendant difficulties also thrown into relief if we consider a point of view from which they are absent: that in which appeal is made to the notion of an explanatory system.

On that view the correspondence between the words spoken by the student and those read by me as I look over his shoulder is explained by his also reading them, so that his inner processes come into play just as mine do. Similarly, that the two of us both see and hear the words he speaks is also part of the explanation; for how else is it to be explained that the words he speaks are the same as the ones I read? Well, perhaps it is a coincidence – except, of course, that this is ruled out by the no-miracles argument. It is worth suggesting, then, even at this early stage of our critique of Wittgenstein, that we should think not of outward criteria

but of a system of reasoning and language use in which inferential and explanatory connections are made between the inner life and overt behaviour of ourselves and other people.

This notion of an explanatory system will be developed later, in the chapter on other minds and elsewhere, my present concern being with the fact that in discussing the concept of reading, I have tried to sidestep questions about the connection between the occurrent and the dispositional, which is why I have yet to discuss reading with understanding. To endeavour to answer those questions is to address what I shall refer to as the intentionality problem, which is freighted with so much difficulty that we shall have to be very careful when steering it in any particular direction. All that I have done so far is to appeal to the everyday fact that in reading aloud I am aware of deriving the spoken from the written word, so that my awareness is occurrent. Since it is also dispositional, the possibility must exist of harmonising the two, which I suggest is a question of analysis, not of dissolving misconceptions.

Return to Wittgenstein on understanding

What I propose, then, is that we set a course for that analysis by ascertaining the extent to which it diverges from Wittgenstein's views. He now returns to his discussion of a pupil learning a sequence of numbers, as with the sequence of even numbers, and he supposes that we have taught our pupil to write the sequence as far as 1000, at which point he disconcerts us by writing 1004, 1008, 1012, whereupon we try to explain to him the error of his ways:

We say: "You were meant to add *two*: look how you began the series!" – He answers: "Yes, isn't it right? I thought that was how I was *meant* to do it." – Or suppose he pointed to the series and said: "But I went on in the same way." – It would now be no use to say: "But can't you see….? " – and repeat the old examples and explanations. – In such a case we might say, perhaps: It comes natural to this person to understand our order with our explanations as *we* should understand the order: "Add 2 up to 1000, 4 up to 2000, 6 up to 3000 and so on." (185)

It is worth pointing out Wittgenstein's arithmetical mistake here, namely that it is not possible if we start at 2000 to arrive at 3000 in increments of 6; therefore, there is no such sequence as the one that he attempts to define. In any case, the point that he wishes to make is perhaps more clearly brought out in the next section:

"What you are saying, then, comes to this: a new insight – intuition – is needed at every step to carry out the order '+ n' correctly." To carry it out correctly! How is it decided what is the right step to take at any particular stage? – "The right step is the one that accords with the order – as it was *meant*." – So when you gave the order + 2 you meant that he was to write 1002 after 1000 – and did you also mean that he should write 1868 after 1866, and 100036 after 100034, and so on – an infinite number of such propositions? – "No: what I meant was, that he should write the next but one number after *every* number that he wrote; and from this all those propositions follow in turn." – But that is just what is in question: what, at any stage,

does follow from that sentence. Or, again, what, at any stage we are to call "being in accord" with that sentence (and with the *mean*-ing you then put into the sentence – whatever that may have consisted in).It would almost be more correct to say, not that an intuition was needed at every stage, but that a new decision was needed at every stage. (186)

He goes on:

"But I already knew, at the time when I gave the order, that he ought to write 1002 after 1000." – Certainly; and you can also say you *meant* it then; only you should not let yourself be misled by the grammar of the words "know" and "mean". For you don't want to say that you thought of the step from 1000 to 1002 at that time – and even if you did think of this step, still you did not think of other ones. When you said "I already knew at the time….." that meant something like: "if I had then been asked what number should be written after 1000, I should have replied '1002'." And that I don't doubt. This assumption is rather of the same kind as: "If he had fallen into the water, then I should have jumped in after him." – Now, what was wrong with your idea? (187)

There is a question here as to what exactly Wittgenstein implies, one answer to which is provided by certain commentators. In a well-known exegetical work G.P.Baker and P.M.S.Hacker take Wittgenstein to be challenging the usual view of what it is to follow a rule and of how it connects with imparting meaning to the expression for the rule. He would say, on their reading, that although the teacher in giving the order meant that the pupil should write 2002 after 2000, this is not to imply a semantic event or process or an act of meaning. The reason is that 'I meant' and 'I knew' refer to dispositions, in this case to the teacher being disposed at the time to give the successor to any term of the sequence. It involves not an act of meaning but mastery of a technique. (1985, p.74)

This mention of technique introduces us to Wittgenstein's thesis equating meaning with use, to be discussed in the next chapter. In the meantime, since one of the arguments thought to support that thesis is that 'I meant' and 'I knew' do not refer to a past event or process, let us consider what happens in an everyday situation. Suppose that at midday I answer a question by saying that Scott is in Spain, followed immediately by my saying that Scott is on the moon, where these are replies to two different individuals asking, 'Where is Scott?' If the possibility of super-luminal travel is discounted, then I must be referring to two different people named Scott, in response to what I understand to be the different references made by my interlocutors, one of whom is an astronaut's husband. But this is to say that at midday I imparted a particular meaning to 'Scott', followed immediately by a change of meaning, which occurred at a particular time, where that occurrence was located in my mind.

This is the conventional account of the temporal status of imparting meaning to an expression, a fact which indicates that we need to clarify what the issues are, especially in view of its being the case, surely, that it is legitimate to speak of dispositions obtaining at a

particular time, at which point the expression we use has a meaning, one that is determined by what we are disposed to say and to do. That said, there is a suggestion, associated with Wittgenstein, that the notion of an act of meaning is misleading if it implies that 'to mean' refers to an event or process that is semantically and epistemically self-contained, in the sense that the meaning we give to an expression is completely encapsulated in it, and that we know it directly, just as we know our own sensations. If meaning is acknowledged to depend on dispositions, then presumably they, too, would have to be taken to be objects of direct knowledge, if the notion of an act of meaning is to make sense.

Since none of this is clear, let us press further into the tunnel in the hope that a light will appear at the end of it. After section 187, the next few sections deal with what it is for mathematical application to be determined by what is meant. We shall return to these, but here is what Wittgenstein then has to say about understanding:

"It is as if we could grasp the whole use of the word in a flash." Like *what* e.g.? – can't the use – in a certain sense– be grasped in a flash? And in *what* sense can it not? – The point is, that it is as if we could 'grasp it in a flash' in yet another and much more direct sense than that. – But have you a model for this? No. It is just that this expression suggests itself to us. As the result of the crossing of different pictures. (191)

Let us return to Baker and Hacker in their capacity as expositors of Wittgenstein. In their commentary on these sections, they take Wittgenstein's interlocutor to be much exercised by the problem of reconciling what seem to be conflicting notions of what it is to understand or to mean. On the one hand it cannot be that we grasp in an instant the whole use of an expression, since that use is extended in time; on the other, it is as if in some sense we do, the use being both present and not present, so that 'what we grasp, when we understand a word, must be something that 'contains' the whole use in some special way, something from which the applications of the expression will unfold automatically, inexorably' (p.112). But this is taken to be incoherent, hence Wittgenstein's claim that we have no model for it. What he himself would say, according to these commentators, is that there is nothing problematic about the notion of grasping the whole use of a word in an instant; it is just that the interlocutor as philosopher seeks the 'kind of understanding of a symbol in which we grasp, as it were, the spirit of the symbol which hovers over all its applications' (p.113). The truth is that:

The life of a sign lies in its use; our immediate understanding of a sign does not consist of our 'laying hold, by means of our logical faculties' of its sense, which is a sort of logical machine, but of our mastery of the technique of its application. (Ibid)

With the help of these elucidatory remarks, we should now have a clearer idea of Wittgenstein's views, except that the commentary itself stands in need of elucidation. What is

one to make of a mistaken notion of understanding that involves the denial that we grasp the use of a word in an instant, and for the extraordinary reason that the use is extended in time? What is one to make of its incongruities when they force us to say that the use of a word must somehow be contained in the act of grasping its meaning? Even Wittgenstein himself has to acknowledge that we have no model for this, that to speak of the use being present, yet not present, is to be under the sway of an expression which imposes itself on us, as if induced by cognitive stress. But if this is to imply that the expression is nonsensical, then how is it connected with the ways in which we are misled?

What we are told is that it points to the conflict in our minds between the notion of understanding as a directly known mental state and as manifested in use or application. If these notions are to be set at variance, then this constrains the ways in which we may apprehend them. As regards understanding being a mental state, it is easy to imagine Wittgenstein agreeing that it is, though with a caveat as to what could be meant if we are not to be misled. Our concern is with what could be meant if we *are* to be misled, or if Wittgenstein's interlocutor is, and it seems to me that we require that understanding be regarded as a mental state to which only categorical truth conditions apply, in which respect it is assimilated to feelings and sensations. This is to exclude dispositions, and the reason is that nothing counts as direct awareness of a disposition, which is in a different category from mental states such as that of feeling tired. It is arguable that the claim to feeling tired is only evidentially, not conceptually, accountable to surrounding facts; these are such that strength of evidence varies with the inductive link between being tired or showing it at a particular time and one's behaviour at that and other times, all within a context in which possible causes of tiredness also play a part. If the claim to understand is in the same way subject to categorical truth conditions, then this is to say that people's cognitive behaviour, for instance in applying a formula, is to be regarded as an evidential basis for crediting them with understanding, which they themselves are directly acquainted with, just as they are with their own tiredness or pain.

It seems that we now have at least some prospect of making intelligible a particular claim: that our muddled notion of what it is to understand evokes in us a picture of the application of the expression being somehow contained in the act of understanding. If this is to make sense in terms of an underlying clash of views, with opposing sides jousting to unseat each other, then its opponent is any view that emphasises the role of dispositions, as with Wittgenstein's thesis equating meaning and understanding with use; or, if different, the everyday notion of the grasp of a rule being manifested in application. Since I shall try to

show, in due course, that Wittgenstein's thesis is untenable, it needs to be emphasised that the view of understanding as being self-contained has other rivals. The one to be singled out is the irreducibility thesis, which I shall introduce when analysing intentionality in terms of a system. The importance of dispositions is recognised in that analysis, as it seems to be by Wittgenstein, so that it is common ground that to understand is to be suitably disposed. Thus, we can avoid pre-judging various issues by the simple expedient of setting the dispositionality requirement against the view that understanding is self-contained.

Suppose, by way of illustration, that I am having a rule explained to me, say the nth place rule for the even number sequence, and that at some point I exclaim, 'Now I understand'. Then Wittgenstein's interlocutor would find himself in a quandary here if he equivocated between the one conception and the other. On the one hand, he would say that I am in an autonomous mental state of understanding; on the other, that to understand the even number sequence is to expand it or to be so disposed, a disposition not being something with which one can have direct acquaintance; and yet I claim to understand the sequence now, and to know that I do. At last, then, we have a genuine conflict of views, or in the case of a single individual, such as the interlocutor, a mutually incompatible set of ideas as to the nature of understanding. Before we enquire into the significance of this antithesis, we need to ask whether it is really true that Wittgenstein concerns himself with it, relevant to which is the following passage:

"But I don't mean that what I do now (in grasping a sense) determines the future use *causally* and as a matter of experience, but that in a *queer* way, the use itself is in some sense present." – But of course it is, 'in *some* sense'! Really the only thing wrong with what you say is the expression "in a queer way". (195)

Hacker and Baker take this to mean that in his muddled way the interlocutor is both attracted and repelled by the thought that the present act or mental state of understanding is causally connected with subsequent application. It would seem, then, that the difficulty faced by the interlocutor is very much that of reconciling views of understanding which are opposed in the way that we have taken them to be. This is borne out by other textual evidence, both in the quoted passages and elsewhere in the *Investigations*.

Despite the importance of such exegetical matters, the use I wish to make of this account of opposing views does not depend on its fidelity to Wittgenstein's meaning, for I maintain that if the quoted passages are to make any sense at all, then it has to be along the lines I have suggested. But now, it surely is obvious that this is a peculiar kind of sense, for what I have taken to be a distorted view of understanding as a mental state with categorical truth conditions is one that no philosopher would find reflected in his thought, so obvious is

the crucial role played by dispositions in intention concepts, one associated with the fact that very often we do *not* know whether we understand whatever it may be. If, however, we reinstate this role, then it is no sooner done than the conflict in question proves hollow, with empty suits of armour falling to the ground.

Return to Wittgenstein on meaning

Let us now examine what Wittgenstein has to say about application being determined by what one means, this being very similar to his treatment of what it is to understand but still deserving of separate consideration.

Here I should first of all like to say: your idea was that that act of meaning the order had in its own way already traversed all those steps: that when you meant it your mind as it were flew ahead and took all the steps before you physically arrived at this or that one.
  Thus you were inclined to use such expressions as: "The steps are *really* already taken, even before I take them in writing or orally or in thought." And it seemed as if they were in some *unique* way predetermined, anticipated – as only the act of meaning can anticipate reality. (188)

"But *are* the steps then *not* determined by the algebraic formula?" – The question contains a mistake.
  We use the expression: "The steps are determined by the formula….." *How* is it used? – We may perhaps refer to the fact that people are brought by their education(training) so to use the formula $y = x^2$, that they all work out the same value for y when they substitute the same number for x.Or we may say: "these people are so trained that they all take the same step at the same point when they receive the order 'add 3' ". We might express this by saying: for these people the order "add 3" completely determines every step from one number to the next.(In contrast with other people who do not know what they are to do on receiving this order, or who react to it with perfect certainty, but each one in a different way.) (189)

He now goes on in the same section to point to another sense of 'determine', which is the sense in which a formula in x and y may uniquely determine a value of y for any given value of x. One example is $y = x^2$, as opposed to $y \neq x^2$ [or $y = \sqrt{x}$]. This is followed by:

It may now be said: "The way the formula is meant determines which steps are to be taken". What is the criterion for the way the formula is meant? It is, for example, the kind of way we always use it, the way we are taught to use it.
We say, for instance, to someone who uses a sign unknown to us: "if by ' $x!2$ ' you mean x^2, then you get *this* value for y , if you mean $2x$, *that* one." – Now ask yourself: how does one *mean* the one thing or the other by " $x!2$ "?
  That will be how meaning it can determine the steps in advance. (190)

To begin with, let us consider section 189. Here as elsewhere it is clear that he takes a third-person point of view, as when he says that the order 'add 3' determines each step for those who respond to it in a similar way, in contrast with those who react to it with certainty, but each in a different way. This is, of course, highly debatable, not to mention wrong, for if

other people all agree in their response to the order by writing 3, 6, 9, 12, 15, 179, then I shall continue to believe that the next term is 18 and that all the terms are completely determined.

This, however, is to cover old ground, so let us move on to Wittgenstein's views on traversing all the steps in advance and their being determined by what is meant. In the quoted passage he discusses various ways in which the steps of a formula are thus determined, the first of which involve training and agreement in use, which I shall consider at a later stage. He then illustrates that sense of 'determine' in which the co-ordinate values of a function are paired off, as, for example, with $y = x^2$, for which each value of x uniquely determines a value of y, though not vice-versa. In these ways, he says, it makes perfect sense to speak of the steps being determined; but he also implies that in the shadows there lurks another sense of 'determine', one that has no real application but to which we are in thrall. He associates it with the picture we have of the steps being traversed in advance, the impossibility of which, when we try to take it literally, is perhaps meant to correspond to inconsistencies in the notion we have of the steps being determined in this way.

The problem, as always, is to try to get a purchase on the ways in which we are supposed to be misled, as was also the case in our previous discussion of understanding. Clearly, these cannot be the same as what for Wittgenstein are the legitimate ways in which one may speak of the steps being determined. This gives us a negative indication of what could be meant, but it is not enough to pinpoint it, this also being true of its being said that for us, or for the interlocutor, it is as if the steps are traversed in advance. The reason is that this is obviously a metaphor, and not one in which the underlying sense is clearly visible. Perhaps, then, we should seek enlightenment in Wittgenstein's own comments in section 188, already quoted:

Thus you were inclined to use such expressions as: 'The steps are *really* already taken, even before I take them in writing or orally or in thought.' And it seemed as if they were in some *unique* way predetermined, anticipated – as only the act of meaning can anticipate reality.

But now, the first sentence is just a paraphrase of its being said that the steps are traversed in advance. If we turn to the second sentence, then a dictionary may be useful here, and mine defines 'predetermined' as 'decided or established in advance'. This is a skeletal definition, as it were, which in a particular case it is left to the context to flesh out. In the present case the reading it would be natural to make is that the steps are predetermined in the sense of being determined by criteria of correctness already established. But since they *are* determined in this way, we would not be wrong to think that they are; therefore, it cannot be what Wittgenstein has in mind when he speaks of our being misled. Since we are no further

forward, let us turn to Wittgenstein's emphasis on the steps being '*uniquely*' predetermined. If this refers to there being only one correct step at each stage, then Wittgenstein himself has already illustrated this as a legitimate use in his example of the function $y = x^2$, a fact which rules it out as what he has in mind. Perhaps, then, the sense of 'uniquely' is given by 'only the act of meaning', so that it refers to the special way in which such acts anticipate reality. Unfortunately, we are still running on the spot, for it is just this special kind of anticipation that we are seeking to have explained.

Perhaps all that can be said, in line with the parallel discussion of understanding, is that the only way to make sense of these sections and the commentary on them is to take the interlocutor to believe that acts of meaning are such as to have categorical truth conditions. This is to portray him as a very idiotic Watson to Wittgenstein's Holmes, one who seems forced into the view that to mean a particular sequence, say, is to consciously represent each term of it in the act of meaning. Anxious to escape this absurdity, he clutches at metaphor and declares that the steps are traversed in advance, really already taken, and so on. Or, perhaps these elements in his reasoning are combined in some other way; but it hardly matters, for on a par with the absurdity just mentioned, as in the case of Wittgenstein on understanding, is the absurdity of setting the interlocutor up to fail in this way, when even the most cursory glance at any intention concept will instantly reveal the crucial importance of dispositions.

Let us now try to anticipate the ways in which the Wittgensteinians would counter the criticisms just made. They would perhaps object that for Wittgenstein such expressions as 'traversing the steps in advance' are simply nonsensical, as indicated by his saying that we have no model for them. If the interlocutor's way of speaking is devoid of sense, then perhaps an objection to our criticisms is that the conflicting views which we have attributed to him are not the source of his confusion, which they therefore do not explain, so that there is no reason to believe that he holds them. Perhaps the idea is that we are not so much misled by the expressions or pictures to which we are in thrall as blinded by them; thus, they need to be eclipsed by correct reasoning if we are to realise the truth of the thesis equating meaning with use. Or again, it would be said that Wittgenstein's aim is to dissolve the difficulties in which we are mired in our confusion as to what it is to mean and to understand.

To this there are two rejoinders. The first is that from the interlocutor talking nonsense nothing follows as to the cogency of Wittgenstein's own thesis, which would have to be separately supported by argument. The fact is, however, that Wittgenstein is not just clearing the ground but preparing it, indeed erecting his theory on it, for in these and other

sections that thesis is taken to be underpinned by his exposure of the interlocutor's muddled thinking. If that unfortunate individual is shown to be incoherent, then we may ask why we should embrace Wittgenstein's views rather than, for instance, the irreducibility thesis – or, there again, why we should not take the entirely different approach of *analysing* intentionality. Secondly, it is in any case misleading of Wittgenstein's commentators to maintain that the interlocutor's expressions are nonsensical, for these philosophers repeat them and paraphrase them as if they have meaning, as does Wittgenstein himself, for instance when he says, again already quoted from section 188: 'Here I should first of all like to say: your idea was that that act of meaning the order had in its own way already traversed all those steps: that when you meant it your mind as it were flew ahead and took all the steps before you physically arrived at that or that one.' As a philosopher one does not need a special interest in the semantics of metaphor to realise that Wittgenstein here imparts meaning to these words.

Perhaps this discussion may be summed up as follows. Wittgenstein wishes to convert us to the belief that we are misguided in our notion of what it is to understand or to mean. With regard to what it is to mean one thing rather than another, he maintains that there are misconceptions in the usual philosophical notion of correct use being determined by what is meant. After illustrating various senses of 'determine' in this connection, he suggests that we mistakenly think that it has another sense, or that there is another sense in which the steps are determined by what we mean; and he also implies that our being misguided in this way is linked to the picture we have of the steps being traversed in advance. He then indicates that such a picture makes no sense, and thereby suggests that we are wrong in the notion we have of the steps being determined. The impression given, as elsewhere in the *Investigations*, is of peering into the depths of truths too profound to be stated clearly, and this owes a great deal to his style of writing. His sentences are on the one hand oracular and opaque, and on the other they announce their importance by the profligate use of italics. The effect on the reader is that he or she is left with a feeling of intellectual unease, a prisoner in Wittgenstein's camp, as it were, who has been made receptive to the new ideology of equating meaning with use.

That for Wittgenstein there are discontinuities in our notion of what it is to mean or to understand is a proposition to which many commentators assent. For instance, in *Meaning, Understanding, and Practice* Barry Stroud's interpretation of Wittgenstein's views on intentionality accords with mine, except that he is much less critical of them. Wittgenstein's target, he says, is the notion of understanding as a mental state, where this is to imply that we should be able to introspect it. What we thereby encounter, however, are non-intentionally

describable mental phenomena, as when we read and understand a sentence, a process which may nevertheless be described as that of looking at marks on paper or screen. Although this experience of understanding a text may be very different from when one has no idea what the sentence means, such a difference cannot be that which separates understanding from the absence of it, and for the simple reason that understanding is intimately connected with application and use. Stroud's Wittgenstein then cautions us against appealing to the existence of an occult cognitive realm in which understanding has its being and generates its own application. Rather, we should identify understanding with application itself, or with use, custom and practice in the public domain.

Irreducibility and the private language argument

My own view, which I share with Colin McGinn and discuss in the next chapter, is that intention concepts are irreducible, which is to say that understanding, for instance, is in a category of its own and cannot be reduced to mental phenomena non-cognitively described. There are places at which Stroud seems to veer towards this position, but he turns away before coming into conflict with Wittgenstein. It seems to me, however, that the equating of meaning and understanding with use does indeed conflict with the irreducibility thesis. Consider the fact that if understanding is in a category of its own , then the cognitive content of, for instance, my reading of a text can be captured only by its being said that I understand what I am reading. It sounds awkward to speak of the exercise of silent reading as being a use of words, let alone a public one, but I understand them only if I am suitably disposed, and this provides a connection with use. I must, for instance, be able to act on any information I have gained from the reading, as also to explain and paraphrase the content of the text, and perhaps to use the words in sentences of my own. But now, all these abilities themselves involve understanding, so that I understand what I read only if I understand other things, this being the connection with dispositions.

There is a crucial point here, as we shall see when we focus on the notion of a system as part of the analysis of intentionality. What matters in the present context is that the irreducibility thesis runs counter to that which equates meaning or understanding with use, since the use in question must itself be informed by understanding. There is, of course, an everyday connection with use, as when we say that tensed sentences are used to refer to the past, present or future, or that understanding is not purely occurrent, as a sensation would be, but is linked to being suitably disposed with regard to application. But this is truistic and it cannot be what Wittgenstein had in mind, even if that is the impression he sometimes

conveys. Rather, he is taken to mean that we are led astray by the notion of understanding as a mental state, or as a mental state belonging to an individual considered in isolation, and that we are brought back into the fold only if we grasp the link between inner processes and outward criteria, itself connected with the impossibility of a private language.

If we now embark on a discussion of the private language argument, then it has to be said at the outset that the argument presents serious difficulties of interpretation; and the claim has been made, indeed, that the literature contains at least five principal ways of interpreting it (Law, 2004). This difficulty is in part connected with the fact that it cannot be a condition of referring to one's own sensations that others happen to understand the reference, for they may not; therefore, what is required is that in principle they could. This is vexingly obscure, not least because much of one's inner life is exhibited to others *only* through the words by which it is described, as with dreams and thoughts and indeed with much of what we experience; and yet we are told that for such description to be meaningful it has to be conformable to outward criteria of correct use.

Wittgenstein's arguments against private language are in concentrated form in section 258, where we are asked to imagine that in a diary one writes a sign, say 'E', whenever one feels a certain sensation, which we may suppose to have no characteristic outward expression. Perhaps one concentrates on the sensation when writing the sign, in order to make it a sign for the sensation. But Wittgenstein claims that in this case there is no criterion of correct use for the sign, so that, 'One would like to say: whatever is going to seem right to me is right. And that only means that here we can't talk about 'right' '.
Could it not be, however, that such criteria are provided by one's memory of the sensation? Rejecting this suggestion, Wittgenstein says that it would be like trying to check one's memory of the departure time of a train by calling to mind an image of the timetable, his point being that this mental image would itself have to be tested for correctness if it were to serve as confirmation of the memory. There are, of course, methods of verification in such cases, but these consist not in conjuring up a mental image but in consulting the timetable or in some other form of interaction with the world.

A rejoinder to this argument is given by Ayer, who maintains that such methods are themselves based on what he refers to as acts of primary recognition, the same as with recognising a sensation, and that all facts and memories depend for their confirmation on memory itself. Referring to Wittgenstein's example, he claims that there is no crucial difference between relying on one's memory of a timetable and actually consulting it, for even when the figures are in black and white one still has to recognise them. Even if these are

checked in their turn, giving rise to further checking, and so on, then, 'unless it is brought to a close at some stage, the whole series counts for nothing. Everything hangs in the air unless there is one item that is straightforwardly identified' (1985, p.76).

This line of reasoning is consonant with our own emphasis on the primacy of the first person point of view, but the role of dispositions has also been emphasised, and this accords with Wittgenstein's argument insofar as there must be more to ostensive definition than writing down a sign in association with observing or experiencing the object it purports to represent.[4] Dispositions supply a missing factor, and this is to say in the present case that the diarist writes 'E', or would write it, every time he has the sensation E. One may imagine, perhaps, that his aim in keeping a record is to explain the occurrence of the sensations. For Wittgenstein, however, it would seem that more again is needed, for he would say that in the absence of public criteria of correct use nothing counts as the diarist correctly or incorrectly identifying the sensation. This returns us, by way of contrast, to Ayer's argument about the fundamental role of primary recognition.

In sympathy with that argument, I now wish to suggest, by borrowing from the anti-sceptical thesis developed in chapter seven, that the taking for granted of memory is a pre-condition of reasoning and language use. This is to say that serious doubt can attach only to particular memories, or kinds of memory, for if it were universal then it would apply even to the memory involved in the process of doubting, which therefore could not get started. Seen in this light, memory is found to be self-intimating; and Wittgenstein's claim that it requires independent criteria of correctness is thereby undermined. In the later chapter I argue that memory belongs within a system in which it is inextricably linked to expectation and prediction. This is not a problem for Wittgenstein's diarist, not even within the confines of noting one's sensations, or not if we suppose that his diary entries include the times at which they occur, and perhaps their duration and intensity, on which basis he becomes aware of patterns of recurrence and is then able to predict the sensations. Besides, there would be innumerable other instances of his linking the past and the future in this way, as when he opens the diary and expects to find previous entries.

---

[4] Wittgenstein is commonly taken to be attempting to dispel the misconceptions arising from our picture of the meaning of a sensation word being fixed ostensively. See, for example, Marie McGinn (1997). This is worth the attempt, for it may be, for instance, that we are misled into assuming that ostension is necessary. Not only is this not the case, but if we keep in mind the distinction between having a concept and learning it, then it is arguable that one can grasp the meaning of a sensation word without being able to have the sensation. See page To clear up that misconception, however, is not at all the same as to show that one cannot as a solitary individual invent a personal sensation language. Better, in my view, to focus on the need for a *system* of language use rather than on the need for public criteria.

What I now suggest, rather than that we continue to engage with Wittgenstein on his own terms, is that we examine some of the assumptions underlying the private language argument. In the article on that argument in which he distinguishes five different interpretations, Stephen Law claims that it is common ground between them that Wittgenstein targets our picture of the mind as an inner mental space in which private objects of consciousness exist. Thus depicted, they are private in that they are inaccessible to other people, therefore unknowable to them, so that only their owner can understand the language by which he refers to them. As Wittgenstein puts it, 'The individual words of this language are to refer to what can only be known to the person speaking; to his immediate private sensations. So another person cannot understand the language.'(243) A possible reading here, if this passage is to make sense, is that Wittgenstein is attacking the thesis, sometimes associated with St Augustine (1), that the meaning of a word is the object for which it stands, so that the meaning of 'pain' in 'I have a pain' is my immediate sensation of pain, which is accessible only to myself, therefore knowable only to myself, so that only I can understand it.[5] In that case, Wittgenstein is clearly justified in attacking it, for it is incoherent to such a degree that we need not trouble ourselves as to what the meaning of the sentence 'I do not have a pain', or 'I did have a pain' could on that view be taken to be.

But now, all that seems to follow is that meaning and object are distinct, this being consistent with the theory, which I shall develop later, that meaning and understanding belong within a cognitive and semantic system driven by the interplay between language, experience and the world. Wittgenstein does, it has to be said, refer to the need for 'stage-setting' (257) if the naming of a sensation is to make sense, and this would seem to be in accord with the idea of a system. But it does not follow, as pointed out earlier, that there need be independent criteria of memory, or that inner processes must have natural outward expression if they are to be understood. With regard to this latter claim, the rejection of Augustine's theory still leaves us with a sense in which subjective experience is private, as illustrated when I say that only I can feel my pain. Thus, if two people are simultaneously in pain, then there are two distinct experiences of a sensation which may be qualitatively, but not numerically, the same. In this respect there is nothing special about sensations in a narrow sense, for if the two people look at a uniformly blue wall, perhaps as a distraction from the pain, the colour experience of each of them is private, even if the one is qualitatively similar to the other.

---

[5] Wittgenstein is often taken not only to be attacking Augustine but also to be in revolt against his own earlier philosophy, that of the Tractatus. See, for example, Marie McGinn (1997).

There is, then, a sense of 'private' which seems perspicuous enough, despite all the difficulties attending the concept of personal identity. With regard to it, one point needing to be made is that the ways in which an individual's experiences become known to others belong within a system in which the natural expression of sensations plays only a small part. Even when there are direct outward criteria of this kind, as with my showing signs of pain without describing it, other people may have limited awareness of its location or severity, or of the kind of pain it is. Perhaps the only way in which these details can be supplied is through my use of language. It is true, as agreed earlier, that it is only by observing other people's behaviour that I can have any conception of them as being conscious like myself; but this is too obvious to be interesting, and if it is all that the thesis linking inner processes with outward criteria amounts to, then there is no reason to dispute it.

What that thesis is usually taken to imply, in the context of the private language argument, is that nothing could count as my referring to my conscious processes in ways that no-one else could even in principle understand. If this is just another way of rejecting the Augustinian theory, then it is unobjectionable, as it is, indeed, if taken to be more general and to mean that there cannot be a language that it is logically impossible for anyone other than the speaker to understand. This, after all, is a popular interpretation of the thrust of Wittgenstein's argument, and with regard to it we may ask how such a language could even be thought to be conceivable. Suppose I am the last surviving member of an ancient tribe adapted to life in subterranean caverns beneath volcanoes, and that I inscribe my diary entries in Vulcanese on tablets of solidified magma which I keep on a granite shelf. I write, for instance, that last night I dreamt of a naked emperor stalking the main vent, or that I woke up with a distinctive sensation in my cheek, hard to describe but unmistakeable, this not being the first time for it to occur. My last entry records my concern about what seems to be a sudden rise in temperature. Suppose, finally, that the volcano erupts and that my diary is flung out and chanced upon by a surface-dweller, perhaps a tourist from Bath. Not being a Vulcanite, she cannot understand it; but what could it *mean*, even, to say that it is logically impossible that she should do so? Whatever it is about me, whether the state of my brain or my being disposed in certain ways, that enables me to read and understand the text, or did enable me before my tragic end, clearly it is logically possible that this person should have the same ability, so that she understands the descriptions both with herself as subject, given suitable adjustments, and as predicated of other people.

Rejecting the private language argument

The point, as before, is that from shared linguistic understanding always being logically possible, nothing follows as to its being necessary. Yet this is a deductive leap that Wittgenstein or his commentators seem to make, and it is not the only source of confusion. We have quoted him as stipulating that a private language refers to the speaker's immediate private sensations, therefore no-one else can understand it. But it is only given a particular notion of linguistic meaning, perhaps the Augustinian one, that it follows that no-one else can understand the language. Once that notion is rejected, we still have the immediate private sensations, shared understanding of which is always logically possible, and if these are taken to be ordinary sensations, private in the innocuous sense examined earlier, the word 'immediate' now redundant, then we do in fact have a shared language with which to refer to and describe them. In the private language argument Wittgenstein is usually taken to be saying that this is possible because sensations have natural expression; but what are his grounds for this claim, given that it is illegitimate to leap from the logical possibility of a shared language to its necessity? Even if it is true, which it is not, that all sensations and other subjective experiences to which reference is possible have natural expressions or some other outward form, why should this be a condition of their being captured by language and of our being able to refer to them? Perhaps it is Wittgenstein himself who assumes that if this condition is not fulfilled, so that there is only the individual's private sensation, then other people could not understand his references to it. He then goes on to claim that understanding would also be impossible for the individual himself, and somewhere along the line a leap is made from the logical possibility of shared understanding to its necessity.

If the assumption really is that others could not understand such references to private objects, then it is plausible to interpret the private language argument as being verificationist, as has often been done, and then to take Wittgenstein as being inclined towards behaviourism, or as attempting to straddle realism about mental content and a behaviouristic approach to conditions of reference. There is, it seems to me, a lack of cogency in Wittgenstein's views on private language, and what is lacking in particular is an argument strong enough to overcome the counter-intuitiveness of the claims that I can understand the putative reference to my own sensations only if other people can, not just as a logical possibility but as a fact about their grasp of concepts, and that they have that ability only if my sensations are such as to be publicly exhibitable in one form or another. I can easily imagine, surely, that other people cease to understand what I say, and without my own understanding being affected. As for private objects needing outward signs, the only

indication to other people that, for instance, I have a mental image of a carpet mite admiring a chimney breast is the reference I make to it. Why should it be thought that a mental image needs to be publicly exhibited, an object of direct knowledge, in order to be captured by language? I cannot observe past events, but this does not affect my grasp of what is said about them.

If in this way we take seriously what common sense would suggest is implausible, and if we do it for insufficient reasons, then perhaps the explanation is partly that we are predisposed, before encountering the private language argument, to accord greater epistemic and even conceptual status to public objects than to private impressions, where to speak in this way would be natural to one who was thus disposed. It marks the contrast between conscious content phenomenally conceived, in which light we take it to be fleeting and insubstantial, and the bodily movements and other outward signs by which it may be manifested, for the human body is a relatively enduring object solidly anchored in the physical world.[6] Such a contrast, or the appearance of it, may then be accentuated if we assume that because one's inner life is inaccessible to others, their knowledge of it must be mediated by direct observation of one's behaviour and relevant circumstances. This, in my view, is an assumption which obscures a great deal that should be clear, including the fact that direct knowledge is much overrated. If my observations indicate that a particular individual is in pain, and if I say that although his pain is inaccessible to me I am able to infer it on the basis of his behaviour, which I directly apprehend, then this is correct in itself but misleading in what it may tempt us to believe. The fact is that I cannot understand his behaviour at all, let alone draw inferences from it, unless I register it as a temporal sequence of events, these consisting of actions, gestures, expressions and so on. But this is to say that the events by which the narrative of his observed behaviour unfolds up to the ever-shifting present moment have been swept into the recent or immediate past, where they are just as inaccessible to me as is his present pain.

Furthermore, if I take his behaviour to be characteristic of pain and to constitute evidence for it, then inferentially I go far beyond my observations into the intricacies of an explanatory system involving my own and other people's subjective experiences and our outward behaviour in the context of our interactions with the physical world. Not only that,

---

[6] For instance, David Pears (1988), in his examination of the private language argument, takes Wittgenstein to mean that because sensations have no physical substance, so that the concept of numerical identity cannot be attached to them, reference to them is possible only if there exist independent criteria of correctness provided by their connections with an enduring physical world. For disagreement about this being Wittgenstein's argument, and for criticism of the argument, apart from the comments on this page, see Stroud (2000).

but if we now consider the privileged epistemic position of the individual concerned, such that he knows that he is in pain, being directly acquainted with it, then again we may call into question what is usually taken for granted. The fact is, after all, that if he claims to know that he is in pain, then he must understand what he says, a condition of which is that he be suitably disposed, for instance to use the word 'pain' correctly, and that his memory of what he says, and of the sensations to which he refers, is correct. No doubt he is entitled to be confident in these ways, but this does not mark a basic epistemic difference between his self-ascription of sensation and his observation-statements or, indeed, his beliefs about other minds.

If this reasoning is along the right lines, then there are no fundamental epistemic distinctions between what I know of my own subjective experiences, of my own and other people's behaviour, and of the contents of their consciousness. It follows that these different areas of knowledge belong within a system, none of them being foundational or having epistemic priority over the others. That being the case, we should no longer be receptive to the basic thrust of the private language argument: that in order to recognise and describe the conscious phenomena of one's mind one requires independent criteria of correctness found only in the public domain. This follows not at all from the logical possibility of other people understanding one's descriptions, and I think it can be shown, quite easily, that they need not understand them.

Counter-example to the argument

Suppose, then, that I am treated by a doctor for not going out enough, the pills having the interesting side effect of washing away colours from what I see, so that my visual screen, as it were, becomes monochrome, showing only black and white. Then I am still able to assure myself that my errors of public application do not betray a lack of understanding of the words for the colours that remain to me. Looking at a particular swan, I say to other people that it is white, which is what I remember, only to be told that it is black, which makes it an Australian swan; and it may be that the swan appears black to me, or that it appears white, but in either case I am able to differentiate between the two.

All the same, if my experience of which objects appear white and which black differs greatly from other people's, then I may give the impression of not even knowing the meaning of 'black' and 'white', let alone that of the words for chromatic colours. We could go even further and devise a hypothetical situation in which this impression was reinforced, for instance by specifying that the medication affected not only my perception of colour but also

the apparent colour stability of physical objects. The colours would be stable for other people, other things being equal, but for myself they would fluctuate between black and white on different occasions. That being the case, those around me who were unaware of my condition would reject my claim to understand the words for these two colours. Those who were aware of it would perhaps say that they had no way of checking whether I understood. If this is correct, then the problem for Wittgenstein is that an individual may give meaning to colour words even when external criteria for their correct use are unfulfilled. A possible objection here is that in this example my colour words derive from the public language of colours, my grasp of which was already established before I took the medication. But this is a contingent fact about colour concept acquisition, the point at issue being that of whether my use of colour words depends on public criteria; and it would, I think, be easy enough to modify the thought experiment so that I acquired the concept in some other way. I conclude, after this long discussion, that it is not the case that words have meaning only if their use is conformable to public criteria of correctness.

Let us now recapitulate the main points of this chapter and prepare for the next. I have tried to show that we are not in thrall, as Wittgenstein seems to suggest that we are, to misleading and conflicting pictures of what it is to mean and to understand. Rather, our concern should be with questions of analysis, and there is no reason to believe that they have to be dissolved instead of answered, or that only the sceptic about meaning and understanding is able to answer them. That, at least, is the thesis I am trying to establish, to which end I propose in the next chapter to critically examine not only the equating of meaning and understanding with use but also the sceptical view of them taken by Kripke. This is all as a preliminary to a modest proposal of my own, already prefigured in this chapter: that it is only within the context of a system that an adequate account can be given of what it is to mean and to understand

My ultimate aim, of course, is to provide a solution to Hume's problem and that of other minds; and in this chapter I have emphasised the first-person view of understanding and the fact that it is not only occurrent but also dispositional. A question now arises as to how we know that we are suitably disposed, for instance with regard to our understanding of a philosophical proposition, with a similar question – to be discussed in the next chapter – about how we know what we believe. The answers will contribute to its being shown that meaning, understanding and belief are part of the structure of the system in which we are conscious of the world and reason about it, where this includes philosophical discourse about the possibility of scepticism. It would advance matters, too, if we could capitalise on the

suggestion, made in this chapter, that understanding informs reasoning; or, put another way, that from a first-person point of view it is only within a cognitive framework that a judgement about understanding can be made in any particular case.

What it should also be possible to show, again on a subjective view, is that acts of recognition, for instance as informing a perceptual event, are intrinsic to awareness of the world and to the use of language. Unlike Ayer in the quoted passage, I would emphasise that recognition requires that one be suitably disposed. It might be thought that it is epistemically prior to inductive inference, if this is based on recognising past instances; but this is not the case if dispositions are involved, a point that the appeal to the notion of a system, rather than to foundationalism, is meant to bring out. And if logic dictates, as it seems to me that it does, that the inductive sceptic cannot stop short of global scepticism, so that he is committed to denying that we recognise anything at all, a also that we understand anything, understanding being dispositional, then it should be possible to show that he refutes himself.

Still, much remains to be done if the theory now emerging is to be brought to fruition, and if an answer to the inductive sceptic and to scepticism about other minds is to be plucked from it. As in this chapter, my aim in the next will be to cultivate a particular account of what it is to mean and to understand and of the way in which inference enters into language use. The method, as before, will be to clear the ground currently occupied by Wittgenstein, and also, in this next chapter, by Kripke. It will then be easier to unearth the hidden connections between language and the world.

Chapter 5

Meaning As Use

Continuing with our critical examination of Wittgenstein's views, let us now return to what he says about use, custom and technique. With regard to this latter, there is considerable overlap between the notion of ability or technique and that of understanding or meaning, the verbal expressions of which are in many cases interchangeable. It matters little whether we say that we understand the even number sequence or that we are able to expand it; this is also true of our saying that we know how to ride a bike or that we are able to ride a bike. It is undeniable, then, that to exhibit understanding is to display ability or technique, though this does invite a question as to the significance of that fact. In seeking an answer, we should perhaps begin by considering physical rather than mental ability, a difference which the following excursion into traffic will illustrate.

Since riding a bike has been mentioned, let us imagine that a robot has been programmed to learn to cycle, a skill which it quickly masters as it becomes adept at weaving between vehicles and ignoring traffic lights on red. There are circumstances in which it would not be misleading to say that the robot knows how to ride a bike, though we should perhaps have to stop short of describing it as being angry and disappointed when it fails its cycling proficiency test. This is all from an external point of view, the robot not being credited with an inner life, and what it shows is that criteria of cycling ability are behavioural, as to some extent are criteria of understanding. The difference, if previous arguments are correct, is that criteria of understanding in others are linked to their being conscious, and in the same way in which we attribute understanding to ourselves. Inescapably, then, we are drawn back to the first-person point of view, from which perspective it is open to me to claim not that understanding *is* technique, whatever that could mean, but that it is acquired through training and exhibited as a technique, or not outwardly exhibited at all, as when I re-read these words or read other text, or register what other people say.

Since it would seem from this account that the notion of technique as applied to understanding is consistent with any analysis of what it is to understand, there may be little to be gained from considering Wittgenstein's views on technique in isolation from the rest of what he has to say. Perhaps we should look to him to explain what else is involved when we acquire a particular technique of understanding; and to that end we shall now enquire into what he has to say about the connection between meaning or understanding and use. He writes:

Now think of the following use of language: I send someone shopping. I give him a slip marked "five red apples". He takes the slip to the shopkeeper, who opens the drawer marked "apples"; then he looks up the word "red" in a table and finds a colour sample opposite it; then he says the series of cardinal numbers – I assume that he knows them by heart – up to the word "five" and for each number he takes an apple of the same colour as the sample out of the drawer. –It is in this and similar ways that one operates with words.– "But how does he know where and how he is to look up the word 'red' and what he is to do with the word "five"? – Well, I assume that he *acts* as I have described. Explanations come to an end somewhere. – But what is the meaning of the word 'five'? – No such thing was in question here, only how the word "five" is used. (1)

Here we can at least catch a glimpse of what Wittgenstein has in mind when he equates meaning with use. The shopkeeper is depicted as having to take a circuitous route to the understanding of English, or of particular instructions, and we have to imagine, I suppose, that he is not able to take the shortcut of translation into a language he already knows. The emphasis is on what he does, and this is meant to show that in a wider context any question as to the nature of meaning can be answered in terms of use. This is not to imply that there is no such thing as meaning, for in English 'meaning' and 'to mean' have a use, the whole of which forms a practice, or, as Wittgenstein would say, a language-game

The problem is to try to give substance to this thesis, and one place to start is with the actions of the shopkeeper. We saw that he carries out the instruction by breaking it down into simple steps and matching words against colours and objects, so that the meaning of the words lies in the use he makes of them. The implication here is that this is also true of actual linguistic understanding and practice, so that, for instance, the shopkeeper would quickly learn to understand English in the way that we all do and would be able to dispense with elaborate reminders and carry out the order without more ado. That is, he would read the order and immediately place five red apples in a bag, handing them to the customer in exchange for payment. This is what would normally happen, and again it illustrates what could be meant by the use that the shopkeeper makes of words. At this level, indeed, it is easy to imagine replacing them with pictorial signs, in the present case a colour picture of five red apples on a slip of paper. But now, this immediately invites a question as to whether more sophisticated functions of language to convey meaning are amenable to the same interpretation in terms of use. Statements about the past or future come to mind, and at some point we shall have to consider them in connection with Wittgenstein's thesis.

In the meantime, I suggest that we consider that thesis from a first-person point of view. How are we to reconcile it with the fact that conscious processes are essential to what it is to mean and to understand? In search of an answer let us imagine that the shopkeeper is myself, so that we may ask what happens when I carry out the instruction. Suppose that I

open the drawer marked 'apples'; then my powers of recognition come into play, not only in my recognising this instance of 'apples' as the same as the one in the instruction but also in numerous other ways, including my recognising the colour of the apples by comparing it with a colour chart. It is also the case, along similar lines, that there are numerous ways in which my expectations are at work, and these include my expecting to find apples in the drawer. Equally undeniable is that I cannot carry out the instruction without remembering it, and also that memory attends all my actions, not to mention the fact that expectation depends on it. Finally, this is all bound up not just with intentionality but with my having particular experiences, for instance the perceptual experience which is distinctive to seeing the word 'apples' as opposed to some other word.

What makes these facts significant is that recognition, expectation and memory enter into meaning and understanding and underpin them, not only in the present example but in all areas of discourse. Indeed, there is a sense in which they transcend them. If, for example, I try to understand a particular rule, checking whether my application of it is correct, then I engage in a process which involves recognition and so on, and in ways that are independent of whether my grasp of the rule is secure. This connects with our previous claim that it is only within a framework of understanding that I am able to check my application of a rule. Here is what Wittgenstein has to say on the topic of recognition:

Asked "Did you recognise your desk when you entered your room this morning?" – I should no doubt say "Certainly!" And yet it would be misleading to say that an act of recognition had taken place. Of course the desk was not strange to me; I was not surprised to see it, as I should have been if another one was standing there, or some unfamiliar kind of object. (602)

No one will say that every time I enter my room, my long-familiar surroundings, there is enacted a recognition of all that I see and have seen hundreds of times before. (603)

It is easy to have a false picture of the processes called "recognising"; as if recognising always consisted in comparing two impressions with one another. It is as if I carried a picture of an object with me and used it to perform an identification of an object as the one represented by the picture. Our memory seems to us to be the agent of such a comparison, by preserving a picture of what has been seen before, or by allowing us to look into the past (as if down a spy-glass). (604)

And it is not so much as if I were comparing the object with a picture set beside it, but as if the object *coincided* with the picture. So I see only one thing, not two. (605)

In these last two sections Wittgenstein does what one has more or less come to expect, which is to suggest that we are misled by a false picture of recognition, a suggestion which then seems to confer on the first two sections a significance they would not otherwise have had. I do not believe that I have this false picture of recognition, and I already know, as we all do, that there is at least some sense in which an act of recognition does not take place when I

encounter a very familiar object; but equally, there is a sense in which it *does* take place, so that we could happily discuss what exactly happens and how best to describe it.

Rejection of meaning as use

A discussion along those lines would, however, be intricate in its details; but the good news is that we may cut through those intricacies by focussing on the use of language or the level of awareness necessary to its use, for we are not really interested in what it is, for instance, for one carpet mite to recognise another. With regard to language, then, if I see a red apple and describe it as such because it is red and an apple, then this is to imply that I recognise its self-similarity from one moment to the next, as exhibited, for instance, in my continuing with the same description or being so disposed, in which case I must also recognise the different instances of the same words. The crucial point here is that one may argue that recognition and similarity are *fundamental* and *irreducible*, this latter in the sense that nothing can be said about them that does not pre-suppose and utilize them, as in this paragraph and the last. The fact is that if I recognise an object which I perceive, then I do so *because* I perceive it, even if there must always be further conditions to be satisfied. Thus, if I am looking at a red object and recognise it as such by labelling it or being so disposed, then I do so because of the particular perceptual quality of my experience, which irreducibly can only be described as the experience of seeing red, if it is on that basis that I recognise the colour of the object[1]. And that, for many purposes with regard to combating Wittgenstein, is all that we need to be able to establish.

For what could it signify to say that the meaning of such terms as 'recognise' and 'similar' resides in the use we make of them, or that to apply them is to engage in a practice, or a practice according to a custom? If we speak of the use of 'recognition', then what we mean is that we use it to mark our recognition of whatever it may be, which is not a felicitous attempt to elucidate the meaning of the term. As regards custom, if this is to imply communal use and the application of public criteria of recognition, my criteria for whether other people recognise particular colours depends on my *own* recognition of them. One way to sum this up would be as follows: an outward criterion stands in need of inner processes.

This is encouraging, so let us look further afield for other cases which we may exploit as counter-examples, this time not restricting ourselves to the areas covered in the *Investigations,* which tends to concern itself with arithmetical rules, perhaps favoured by

---

[1] It need not be on that basis, for it could be that in poor light I recognise the shape of an object which I already know to be red.

Wittgenstein because it is not clear whether and in what sense the statements of arithmetic refer to anything outside the arithmetical system itself. No such unclarity obscures statements about the physical world, many of which are conveyed by sentences that are tensed, the action set in the past, the present or the future.

If we begin with statements about the past, then we may ask whether their meaning lies in their use. In seeking an answer, our first step, however odd this may seem, is to remind ourselves of the reality of past events, to which end we may conduct an experiment. Place a clock and a Ming vase on the worktop in your kitchen. At midday push the vase over the edge and note that it smashes into pieces on the ceramic tiles and raises a cloud of ashes, the last remains of your cremated parents. The time is now 12.01 and the result of the experiment is conclusive: a minute ago you *really did* smash that vase; also, your parents were once alive and *really did* exist, as did your childhood. Furthermore, although your spouse is out at the moment, you know that he or she also exists and is going to be annoyed and then start worrying about your mental state, all of which you infer from past events. To genuinely doubt such facts about the past is to qualify as being prone to illogicality, since doubt entails knowledge of immediately past events, at least if it is cast in the form of a sentence.[2]

It has been argued that those who equate statements about the past with their use are not denying the reality of past events; but it needs to be asked what they could mean, apart from the entirely uninteresting fact that such statements are used to refer to past events. For instance, we have quoted Wittgenstein as saying, 'But what is the meaning of the word 'five'? – No such thing was in question here, only how the word 'five' is used.' Now try to imagine a similar claim about the statement that a moment ago I placed five apples in a bag.

The fact is that in this connection the thesis equating meaning with use either reduces to the truism that the concept of the past has application or it calls into question the reality of past events, in which case it must be wrong, for each of us *knows* that events in his or her history *were* real, just as present events *are* real. This is so clearly correct that I do not think we need concern ourselves unduly with the appeal that the Wittgensteinians sometimes make to the distinction between truth- and assertibility-conditions. The thrust of the argument is that the correctness of a proposition depends not on truth but on assertibility, itself connected with criteria of correct use, as with ascribing pain to another person. It is very hard to see, however, how such a thesis about other people could stop short of behaviourism, itself hard to distinguish from solipsism. Well, it is possible that I am alone, and on a good day I can

---

[2] See the next chapter.

conceive of circumstances in which I might be led to believe that I am. Where I fail, however, is in trying to make sense of its being said that a proposition about a past event is correct if it is assertible rather than true. Assertibility conditions cannot even be stated without presupposing the past; and a proposition claiming that an event occurred is correct only if it did occur, and that is all there is to it.

Let us now subject the equating of meaning with use to the further test of applying it to mathematical reasoning. Suppose that I try to work out an nth term formula for the finite sequence 1,5,11,19,29,41. Looking at these terms, I would guess – perhaps from experience – that a suitable formula would be quadratic. So let us try $n^2 + n + 1$. No, the first term is wrong, so make it $n^2 - n + 1$. No, this time the second term is wrong, so try $n^2 + n - 1$. Yes, it works for each term. Clearly, this process of reasoning is heuristic, my discoveries determining the inferences that I make. If I find that for $n = 2$, $n^2 - n + 1 = 3$, whereas the second term of the sequence is 5, then the formula $a_n = n^2 - n + 1$ *cannot* be correct. But now, if this reasoning is guided by deductive inference, so that it involves mathematical relations between signs, then what could it mean to say that such inferences are a matter of use or practice? If this is to deprive them of their logical force, then instead of explaining them it eliminates them; and if it is not, then it says nothing at all about them.

Critique of Kripke on rule-following

That, at least, is what common sense would suggest, and if this is to be our guide then the foregoing criticisms have general application to the thesis equating meaning with use. Another approach to them is to ask whether the thesis implies that the use of an expression is unconstrained except by habit or custom. An affirmative answer would again be counter-intuitive, regardless of whether the use in question is that of the individual or the community; and it is not clear that Wittgenstein would subscribe to it. Nonetheless, it is an answer given by some commentators, notably Kripke, whose radical interpretation of the *Investigations* must now be considered. He is of the view, repudiated by Baker and Hacker (1984), that for Wittgenstein the character of rule-following is so extremely paradoxical as to compel the sceptical conclusion that meaning and understanding are non-factual. Setting aside the question of whether it is faithful to Wittgenstein's views, the sceptical thesis propounded by Kripke has been taken to merit examination on its own account, and the place to start is with his argument about what it is to give meaning to an arithmetical sign. Suppose that I denote by the symbol '+' the addition function, the sum of only two terms being considered, and that

I have never summed any terms greater than 56, apart from the present sum, which takes the form 57+68=125. Then we may imagine a sceptic asking how I know that in my previous use of '+' I meant addition, such that this past meaning determines the answer 125. Perhaps I meant 'quaddition' such that $x + y = 5$ if $x, y > 56$. The intention here is not to call into question my memory; rather, it is to cast doubt on the whole idea of what it is to follow a rule or give meaning to a word.

One way to try to meet this challenge would be to appeal to some inner experience accompanying my previous use of 'plus'; but it is common ground that neither meaning nor any other form of intentionality is to be assimilated to sensation, despite such well-known phenomena as flashes of understanding and the sudden revealing of a word on the tip of one's tongue. After all, how could an inner state, an immediate item of consciousness, determine the future application of a rule?

At this point we may try to answer the sceptic by turning to the dispositional element in rule-following. For instance, if I have grasped the addition rule then I am disposed to give to each addition sum a particular unique answer. But now Kripke in the role of sceptic argues that rule-following cannot be equated with an individual's actual dispositions, for then it would not be possible to distinguish between correct and incorrect applications of a rule. The reason is that there cannot be a rule without standards of correctness, so that the concept of a rule is normative. Besides, dispositions are always finite, whereas x and y may take any of an unlimited range of values. Since we appear to be unable to answer the sceptic, the only conclusion, according to Kripke, is that 'there can be no such thing as meaning anything by any word' (1982, p.55).

He now turns to what he calls the sceptical solution to the paradox, by which he means a solution that acknowledges that there is no fact of the matter as to what it is for an individual to follow a rule. What is wrong, says Kripke, is the assumption that an individual considered in isolation may follow a rule, the truth being that it is possible only if one is regarded as a member of a community, except that in a sense it is still not possible, for the rule is constituted by communal practice itself rather than the practice complying with a separately identifiable rule.

If rule-following is thus constituted, then what the sceptical solution entails is that *any* communal practice counts as following a rule, except that this would vitiate the very idea of rule-governed action or meaningful discourse. Indeed, even the expression 'communal practice' would be devoid of content, since no individual could mean anything by it, and a

community is composed of individuals. If this is correct, then there must be flaws in the arguments that lead to the sceptical solution, and in the very notion of a sceptical problem about what it is for words to have meaning. What I propose, then, by way of a preliminary report, is that I outline the case against the sceptic, or my version of it, and then go on to examine the issues in greater depth.

Let me begin with the observation that in his role as a sceptic Kripke has a reason for targeting previous applications of a rule. His reason is that the sceptical problem can be formulated only if one's present practice is not itself open to question. For instance, if I now say that 57+68 =125, the sceptic will agree that by '+' I now mean addition, and that all my words and his have their usual meaning. This has to be conceded, says Kripke, but he adds that once the problem has been stated it poses a difficulty for one's present use of words considered at a later stage. In the light of this concession, I shall now try to show that Kripke's arguments are incoherent, and one reason is that the sceptic's words are no sooner uttered than they recede into the past. The fact that it is the immediate past makes no difference – the sceptic still has to question what was meant by them, that question having to be challenged in its turn, and so on into an endless regress.

A related argument is that it is a necessary condition of language use that words retain their meaning from one instance to the next. Thus, if I say to the sceptic, 'you seemed to agree just now that 57+68=125, so would you still agree that it does?' then the sentence is meaningful only if the repeated items are taken to have a continuing sense, in this case the words 'you' and 'agree', and only if the same is true of any repetition implied in a wider context in which the sentence has meaning; for instance, the context in which 'it does' is expanded to mean '57+68=125'. Moreover, if a sceptic uses a word in a sentence, so that he has to agree that it means the same upon repetition, then the same applies to his use of the sentence, and therefore to his use of the sentences by which he asserts that reference to meaning is non-factual. It follows that he is open to the charge of self-refutation.

Finally, the concept of the past is essential to Kripke's arguments; but if there is no fact of the matter as to what we mean, then how can it be possible to refer to previous linguistic use? Kripke's answer, no doubt, would be that semantic non-factualism is not to be equated with nothing being meant by what we say, for its meaning resides in public practice. But we have seen that the appeal to such practice is highly problematic, so that we are entitled to ask how the distinction in question is to be made. If the sceptic concedes that it cannot be made, he immediately lapses into incoherence, and again the indication is that semantic scepticism is self-refuting, so that non-factualism about meaning is indefensible.

The irreducibility thesis

That being the case, we are left with the common-sense view that facts exist, and in particular that there are facts about what it is to give meaning to a word, such that ascriptions of meaning, as when we say that by 'red' we mean *red*, are truth conditional, as is the claim that we understand colour concepts. This is the thesis of semantic realism, and in favouring it we cannot just ignore the arguments of the sceptic but instead have to give an account of meaning that answers them at the same time as it accords with our intuitions as to what it is to mean and to understand. One possibility would be to naturalise meaning by giving a reductive account in terms of mental and behavioural dispositions; and another would be to develop the thesis that meaning and other forms of intentionality are irreducible, where this does not preclude their also being dispositional. In the previous chapter it was, after all, taken for granted that a condition of understanding a word or sentence is that one should be suitably disposed.

One advocate of the irreducibility thesis is Colin McGinn, who argues that if the candidates for meaning examined by Kripke all fail to qualify, then what follows is not that there is no fact of the matter with regard to it but that it stands on its own, a thing in itself, in the sense of not being reducible to anything else. And the same would apply, of course, to understanding, and no doubt to other intention concepts. Thus it is that for McGinn the crucial question is that of:

why we cannot give the truth conditions of 'he means addition by "+" ' simply by *re-using* that sentence, frankly admitting that no other specification of truth conditions is available – precisely because semantic statements cannot be *reduced* to non-semantic ones. (1987, p.151)

What I now propose is first of all to argue indirectly for this thesis by eliminating one of its competitors, namely that form of naturalised semantics which reduces what it is to mean and to understand to dispositions to behave in ways that satisfy criteria of meaning and understanding. Kripke would contend, of course, that the task of elimination has already been carried out by the semantic sceptic, but my present concern is to show that in any case the view in question is counter-intuitive and clashes with common sense.

If we start with behavioural dispositions, then, and compare them with dispositional properties elsewhere, for instance with glass being brittle, then we find that glass is brittle if it is disposed to break, but that its breaking is itself an event or process, reference to which is indispensable to the concept of brittleness. In terms of disposition and behaviour, to understand is to be disposed to behave in certain ways and to actually behave in those ways while being so disposed. This formulation, in which mental processes are not included,

captures some of the conditions of understanding, at least if we treat it at the level of common sense; but if we see it as part of a reductionist analysis, then difficulties arise as to what could be meant and as to the adequacy of the account it provides.

In order to probe these difficulties, let us first of all more fully present the reductionist case, which is seen to best advantage from a third-person point of view, that in which understanding is attributed to other people. Suppose, for instance, that I judge that another person understands the even number sequence, given his actions and what they indicate about those to which he is disposed. Then it could be argued that his understanding is constituted by his behaviour, not by any inner states and processes which it manifests and reveals. This is not to deny that to credit him with understanding is to imply the existence of conscious processes guiding his actions, otherwise my computer would understand the game of nought and crosses. The same is true of other animals, or higher animals, to which we attribute understanding, for instance dogs and bats, even if the grasp of arithmetical sequences would seem to have eluded them thus far. Human beings are, however, notoriously anthropomorphic, even if they stop short of similarly characterising other species, and it is arguable that we have no genuine insight into the cognitive processes of a dog fetching its lead or of a bat hanging very cleverly upside down[3]. All that we can say, or so the theory goes, is that understanding, even if it requires consciousness, consists in overt behaviour, particular inner processes being unknowable in the case of other animals and inessential in the case of people.

Now it has to be admitted that such a theory may seem plausible, at least until its full implications are brought out and a comparison made, as we have already done, with a first-person view of what it is to understand. If previous arguments are correct, then it is very obvious that the facts of my own understanding are inimical to any analysis that disregards or sidelines conscious processes of cognition, including those that enter into my judgement of another person's understanding. Taking this for granted, let us now probe more deeply than before into the ways in which understanding enters into reasoning. To that end, we should perhaps consider what it is for me to grasp the even number sequence or determine whether I understand it, which I do by reasoning about it or, if reading and writing are involved, by merging together thought, perception and action to yield the inferential steps by which I proceed. Characteristic of engaging with the sequence in this way is that I extend particular segments of it, work out the harder terms or successors in its higher reaches, check that I am

---

[3] Thomas Nagel (1979) asks what it is like to be a bat.

right and perhaps correct any slips or errors, test my dispositions by sampling, and so on, all of which are instances of cognitive processes. Even if I imagine struggling to apply the rule, perhaps becoming aware that my grasp of it is tenuous, then this could occur only within a wider understanding on my part, without which it would make no sense to describe myself as checking my results, detecting my mistakes, and in general trying and failing to understand the sequence. Thus it is that the notion of understanding is intrinsic to cognitive processes, both to engaging in them and to reflecting on them, which itself is a form of reasoning.

What this indicates is that we go wrong if we consider in isolation the grasp of a particular rule, as with the even number sequence. If we check another person's claim to understand it, and if we find from his overt behaviour that he does not, then we might argue that his cognitive processes, whatever their particular form, are compatible with his not understanding the sequence, the grasp of which therefore consists in correct public practice. And if this is the case with any particular rule, then why should it not hold for understanding in general? We have seen, however, that this line of reasoning cannot be correct; this is to say that no analysis that disregards conscious processes can capture what it is to understand, and none that recognises them can be reductive in terms of behavioural dispositions. Even if such processes are given recognition, however, it still needs to be asked whether an account in terms of them could itself be reductive, a question which connects at several points with the discussion thus far.

Perhaps we may answer it by recurring to McGinn's example of the criteria by which it is judged that another person means addition by '+'. It may be, for instance, that my judgment to that effect is based on his answers to questions of the form $a + b = ?$ in which case I myself mean addition by '+', a fact which is essentially linked to the conscious processes by which I reason arithmetically and interpret his responses. But now, it surely is obvious that my understanding of the addition function is not constituted by mental processes or dispositions of mine that are susceptible of a non-semantic description and in that way described. This becomes even clearer the more careful we are not to insinuate semantic presuppositions into, for instance, our view of the mental images which appear to convey our thoughts. With my eyes closed, I can work out the sum of, say, 68 and 57, which in my case involves visualisation; but if in the guise of reductionist I cite the fact that in meaning the numbers 68 and 57 and the addition function, I may conjure up mental images of them, then should I try to 'see' them as imagined lines and curves and thus describe them? But what would be the point? When they play a part in the visualisation involved in my arithmetical reasoning, I do not 'see' them in that way, and if I did then they would have no part to play in

my giving meaning to a sign. To be accurate in my characterisation of them I have to say, or so it would seem, that I have mental images of the numbers 68 and 57 and of the addition function.

Hattiangadi's objections

What we are left with, then, is the irreducibility thesis, and in order to secure it we are going to have to defend it against its critics, this being the aim of the present section. If we start with McGinn's version of that thesis, his views have been criticised by some commentators, including Hattiangadi (2007), whose disagreement with him is part of her rejection of non-reductionism insofar as it purports to meet the challenge of the semantic sceptic. Let us now consider her general position in order to provide a context to her arguments against irreducibility.

As a proponent of semantic realism she makes the case, in my view convincingly, that non-factualism about meaning and understanding is incoherent and self-refuting, as is Kripke's attempt at damage limitation in the form of the sceptical solution. Clearly, Kripke's sceptical argument is deeply flawed, and what is wrong, she says, is his assumption that meaning is normative in the sense of being prescriptive. Rather, it is norm-relative, which is to say that there are standards of linguistic correctness, but without it being implied that when I use a word I ought, in some normative sense, to use it correctly, this requirement being intrinsic to what it is to give meaning to a word. This is not a perspicuous distinction, but the main thrust of her argument, in any case, is that a fully general semantic scepticism is possible only if meaning is taken to be normative, whereas in fact it is norm-relative. The difference this makes, in her view, is that although no attempt to refute the sceptic has yet succeeded, his arguments are not such as to rule out *a priori* the possibility of future success.

Wishing to demonstrate the lack of success thus far, she criticises all attempts to naturalise meaning, and she then targets the non-reductionists, starting with McGinn, who, she says, conceives of meaning and understanding in terms of real and irreducibly mental capacities. Thus characterised, they are essentially the same as dispositions of the kind invoked by Martin and Heil (1998) in their rejoinder to the sceptic. Like McGinn, they treat intentionality dispositions as being irreducibly mental; and, too, they take them to be real in the sense that they are not defined by whichever conditionals turn out to be the case, where this would mean that no distinction could be made between what I am disposed to do and what I actually do. Kripke, it seems, does not regard them as being real in this sense, and that is why he maintains that the dispositionalist cannot account for the possibility of error, since

whatever I do I was disposed to do. The fact is, according to Martin and Heil, that there may be all sorts of reasons why one's actions may conflict with a particular disposition, perhaps because of tiredness or inattention, such actions being separate from those by which the disposition is manifested, so that they are not part of its extension. Thus it is, according to these writers, that dispositional theory is able to accommodate the possibility of error.

Hattiangadi, however, does not agree, and she continues to maintain that the possibility of error is a stumbling block for the dispositionalist:

> I might have several (real) dispositions, including the disposition to add, in some conditions, and the disposition to make mistakes in others. The dispositionalist needs to supply a non-circular distinction between those dispositions that constitute meaning and those that produce errors. Treating dispositions as real does not supply the relevant distinction. (p.153)

What I now propose is that we make a start at resolving the various issues and then resume that task after further examination of Hattiangadi's critique of non-reductionism. Let it be said first that the issues are not conducive to clear thinking, and are so confused, indeed, that even Hattiangadi is not quite correct when she takes McGinn to regard capacities as irreducibly mental (Hattiangadi, p.152). In fact, he contrasts capacity theory with the irreducibility thesis (pp.164-168); but then, he also contrasts this latter with what he refers to as the Gricean programme of semantic reduction (Grice, 1957), despite saying that Grice's analysis of meaning helps itself to 'the notions of intension and belief with determinate conceptual content' (McGinn, p.167). He suggests, indeed, that Kripke fails to distinguish between meaning and concept, a relevant difference being that a concept and its content cannot be treated separately in the way that a word and its meaning can. Thus, the sceptic can require me to show what it is in my present use of a word that confers on it the same meaning as previously, but my present grasp of a concept does not depend on its having the same content as before, for if it has a different content then it is simply a different concept (McGinn, p.147). Could it be, one might ask, given this neglect of an important difference, that other contributors to the debate have also overlooked it and that this helps to explain the insouciance with which they *use* concepts in the very process of purporting to show that their verbal expression has no meaning? We take seriously, after all, the assertion that nothing counts as our use of 'plus' to mean *plus* or *addition*, and we thereby *use* the concept of addition and cannot deny that we understand it.

There is more than a hint of self-refutation here, and what has also been pointed to is a lack of clarity with regard to what exactly one is entitled to assume in debating semantic scepticism without begging the question. One could spend many a happy lifetime

disentangling the various intricacies, but with regard to Hattiangadi's views on irreducibility the matter is more clear-cut than it would be on first encounter with the sceptic, for having already argued for the conclusion that semantic non-factualism is incoherent, she is within her rights to take semantic realism for granted. That being the case, one must ask whether the different strands of her reasoning are fully joined up, for she still insists on trying to show that the irreducibility thesis begs the question against meaning scepticism or evades the issues with which it is entangled.

It would be more reasonable, surely, to argue that once the semantic sceptic has been shown to refute himself, so that the reality of meaning and understanding is restored, then the only task which remains is that of elucidating the semantic and the cognitive and analysing their conditions of application, rather than continuing to grapple with an already defeated opponent. We should, for instance, cease to defend particular forms of dispositionalism against the sceptic and focus instead on the adequacy of the accounts they give of what it is to mean and to understand. This was my project earlier when I rejected any account in terms of naturalistic dispositional theory, notwithstanding the fact, as I see it, that dispositions are an essential component of intentionality. If that theory is rejected, along with other forms of naturalised semantics, the last candidate standing is the irreducibility thesis, which, once appointed, frees us up to investigate intentionality rather than to defend it. Hattiangadi would not agree, for, as quoted, she still requires the dispositionalist to 'supply a non-circular distinction between those dispositions that constitute meaning and those that produce errors'. But McGinn is a non-reductionist, and it is open to him to argue that with the sceptic out of the way, one's possession of the concept of addition can no longer be challenged, the only question being that of what it is to mean *addition* by 'plus'. This, in my view, calls for an answer, not a defence, but in any case it is obvious what one's response should be. I mean *addition* by 'plus' if I am disposed to use 'plus' in my application of the concept of addition, other things being equal, where the ceteris paribus clause allows for the possibility of error. Since Hattiangadi does not respond in that way, it must be that she does not accept the non-reductionist thesis; and we shall have to see whether she makes a case for rejecting it. The point I am trying to make is that one who has refuted the sceptic and established the truth of semantic realism is thereby entitled to adopt the irreducibility thesis, given that naturalised semantics is untenable, and to *use* the concepts formerly in quarantine, in this case that of addition, in response to a question about the use of language on different occasions to convey the same meaning, or about the semantic connection between word and concept.

As I said, Hattiangadi would not agree; and in fact she is not alone in this. We shall see presently that Wright, himself a non-reductionist, complains that McGinn makes no attempt to account for the paradoxical nature of avowals of meaning, the paradox arising from the fact that they combine authoritativeness with being governed by correctness conditions, so that correct usage goes beyond the particular occasions of use on which avowals are made. Miller (1998, ch.6) concurs with this criticism, and he maintains that no theory of meaning can be satisfactory unless it is able to resolve the paradox. This harks back to Wittgenstein on traversing the steps in advance; but, as I said at the time, those who find the everyday notion of intentionality problematic rather than interestingly in need of analysis need to make out a case that does not depend on metaphor in the way that Wittgenstein's does. The question of avowals is interesting, and I shall try to show that it may be answered in terms of a system; but in the meantime let us return to Hattiangadi, who now discusses Wright's treatment of the paradox, his aim being to turn non-reductionism into a satisfactory theory of meaning. According to Wright (1984), the notion of intention should be taken as primitive, avowals of intention being authoritative because my intentions are constituted by what I believe them to be, other things being equal. Thus, if my intention is to use 'green' to mean *green*, then there is no difficulty about how I know what I intend, and the sceptic cannot ask what constitutes my intention. But, says Hattiangadi, 'The sceptic casts a general suspicion over the possibility that we might give an account of any intentional content, including beliefs about our own intentions... Thus, Wright's account presupposes precisely what it seeks to explain.'(p.160) This is to say that she does not accept that intentions are irreducible, and again one would expect her arguments to be other than those of the sceptic, given that she has already discredited him.

I think that enough has been said to make the point about restoration of the status quo after refuting the sceptic, and in fact I had a dual aim in pursuing Hattiangadi into her discussion of Wright, for I wished to introduce the topic of avowals prior to making full acquaintance with it at a later stage. That dual purpose will also be served if we now turn to what she has to say about the non-reductionism of John McDowell (1998), who challenges what he refers to as Kripke's master thesis that mental representation can be separated into vehicle and semantic content, so that we can always ask how a mental image or phenomenally described thought is to be interpreted. As we have seen, Kripke maintains that nothing in my use of 'plus', including any attendant mental objects, constitutes my meaning *plus*. Thus, if the sentence '67 plus 58 equals 125' is projected onto my mental screen, then it has to be interpreted, according to Kripke, and there is no determinate reading of it as

expressing the thought *that 67 plus 58 equals 125*. McDowell's own example is of a person having the thought *that people are talking about me in the next room*, his point being that we should not first of all imagine the corresponding sentence in that person's mind and then ask how it is to be interpreted, thereby allowing the sceptic to argue that there is no fixed interpretation, that any interpretation would need to be interpreted in its turn, and so on into a regress. Once we have guarded against scepticism by taking thoughts to be intrinsically meaningful, then there is no longer a problem about the normativity of meaning and the way is open for the irreducibility thesis to be brought into the breach (p.272).

In responding to this, Hattiangadi does not deny that vehicle and content cannot be separated in mental representations, as with the belief that other people are talking about me in the next room, but she insists that the sceptic can still ask what constitutes such a belief, given that belief implies meaning and that the sceptical challenge is to explain what it is to mean one thing rather than another. As before, it seems to me that the question about belief should no longer be taken to be sceptical, the challenge having been met, but instead should be interpreted in terms of a neutral investigation into intentionality. To make this as clear as possible, let us revisit the start of this discussion and suppose that I announce that by 'plus' I mean *plus*, as in 2+3=5. Much to my surprise, a sceptic, hackles raised, pointedly asks what constitutes my meaning *plus* rather than any of an infinite number of other arithmetical functions that happen to be co-extensive with *plus* in this particular case or for a limited range of arguments. I reply that I have refuted scepticism about meaning, thereby establishing the truth of semantic realism, and ruled out naturalised semantics, for non-sceptical reasons, leaving only non-reductionism in the field. I then continue, 'So, my answer is that what makes it the case that I mean *plus* by 'plus' is that I mean *plus* by 'plus'. The sceptic now retorts, 'But what makes it the case that you mean *case* rather than *quase*? Or *mean* rather than *quean*?' Resisting various temptations, I calmly reply that this exchange between us is proof in itself of the incoherence of the sceptical approach to meaning and understanding, since clearly the sceptic thinks he knows what he means in his contribution to that exchange. I would add, too, that Kripke's device of querying not present semantic content but only sameness of meaning in relation to past linguistic use does not lessen that incoherence, for – as pointed out earlier in this chapter – it is a condition of meaning and understanding that it must always be possible to refer to a statement once made, as in the present discussion between the sceptic and myself. What has also been shown, I trust, is that I do not beg the question against the sceptic in the answer I give him, and that neither does McDowell when

he maintains that vehicle and content cannot be separated, and that what constitutes my thinking that people in the next room are talking about me is just that that is what I think[4].

Intentionality and avowals

The result of all this argumentation is that we are now in a position to take the truth of the non-reductionist and semantic realist thesis for granted when attempting to elucidate the concept of intentionality, which I propose to do by employing the notion of a system. Given this reinstatement of semantic and cognitive facts, it is possible to say a great deal about them, despite the irreducibility of meaning and understanding. The non-reductionist thesis is, in a sense, negative in content, for what it states is just that the semantic and the cognitive cannot be captured by the use of non-intentional language, the role of analysis being to anatomise their occurrent and dispositional aspects and the links between them. It would be foolish, however, to completely ignore the sceptic at this stage, for he exploits the genuinely puzzling nature of intentionality, for instance of what it is to mean and to understand, and of the authority with which we self-ascribe intentional states. If we are to throw light on these matters, then we have to direct it in such a way as to place the sceptic's arguments in permanent shadow. In what follows, I shall try to show when difficulties are encountered that they are not such as to give sustenance to the sceptic. This is not to say that we should seek to dispel all puzzlement, for it seems to me that meaning and understanding are intrinsically mysterious, *pace* Kripke, the point being that if it is an inherent mystery then we simply have to accept it as such and not allow the sceptic to make propaganda from it. Hoping to find enlightenment, then, let us begin with the epistemic status of avowals in statements of intentionality.

As alluded to earlier, in his comments on McGinn's non-reductionism Crispin Wright takes issue with McGinn accepting, without attempting to explain it, the authoritativeness of self-ascriptions of belief, intention, expectation, and so on, and of the claim that we know what we mean. According to Wright:

> Surely, the relevant feature of the concepts in question – the combination of first-person avowability with disposition-like connections to behaviour in circumstances which the avower need not have envisaged – is no sooner marked than anyone of genuine philosophical curiosity will feel his intellectual conscience pricked. How is it

---

[4]Although it is because the sceptic has already been refuted that the answer is not question begging, it is possible to give another reason. According to Wright (1984), the semantic sceptic can be given pause if we compare meaning to intention, for if my knowing what I intend is non-inferential, intention being sui generis, then I can simply remember my past intentions. On this model, if meaning is sui generis in the same way, so that I know what I mean, then I can claim to remember what I meant by 'plus, for instance, and can avow that it is the same as I now mean by 'plus'.

possible to be, for the most part, effortlessly and reliably authoritative about, say, one's intentions if the identity of an intention is fugitive when sought in occurrent consciousness, as McGinn grants that Kripke's Sceptic has shown, and the having of an intention is thought of as a disposition-like state? (2002, p.113)

Wright has his own views on the nature of intentional states, as we have seen, and his concern is with the distinction between self-ascription of experiential phenomena on the one hand and of particular forms of intentionality on the other. Examples of such phenomena are given by 'pain, tickles, the experience of a red after-image, and ringing in the ears' (Ibid). They are characterised by our being aware of them because we have them: 'The subject is authoritative about such states because, since they are events in his consciousness, he is in the nature of the case conscious of them.' (Ibid) If, on the other hand, we take intentionality to be irreducible, so that meaning addition by 'plus' is not to be identified with its mental accompaniments or with any other experiential state, then how can we be authoritative with regard to it? To mean addition involves, after all, having certain dispositions, so that its content goes beyond the present experience. But this marks it out as being epistemically vulnerable in a way that the reporting of an experiential state is not.

What I shall now try to show is that there is indeed a difficulty here, in the sense that there is a question with no obvious answer, but that it cannot be used against the irreducibility thesis and does not lend itself to scepticism about intentionality. Rather, it is a difficulty of analysis, and its resolution depends on the notion of a system in which intention concepts interact with the world. That, at any rate, is what I shall try to show, beginning with its being a mistake to assume, as in the passage just quoted, that intentionality is not to be found in 'occurrent consciousness'. With regard to meaning and understanding I have argued that if a mental image is representational, then it follows from the irreducibility thesis that this aspect of it cannot be captured by phenomenological description. Similarly, I have claimed that descriptions of seeing lines and squiggles cannot do justice to what it is to read and understand a text. It is in this sense that the reader's understanding is irreducible, but clearly it cannot be separated from the process of reading as a perceptual flow, the cognitive experience of which is just as much a conscious state as is ringing in the ears or feeling pain. Thus it is that in such cases one's understanding is occurrent, though in all cases one has to be suitably disposed, which is to say that when understanding is occurrent, as opposed to being purely dispositional, it is always within a dispositional framework. As for what it is for my understanding to be purely dispositional, I understand the ten times table, for instance, even when I am not thinking about it.

This mention of dispositions leads to my second point, which concerns the fact that in maintaining that descriptions of experiential states are categorical, unlike references to understanding, Wright implies that the difference this makes is epistemologically significant. If his reservations about McGinn's approach are to make sense, then it would seem that he takes us to enjoy privileged access to our own experiential states, on the basis of which we are authoritative in our claims to have them. Such claims are more secure than the corresponding claims about belief, expectation and other intentional states, which obviously have a dispositional component. It seems to me, however, that there is no such difference, for to self-ascribe a conscious experience is to bring into play one's *awareness*, for example of ringing in one's ears, which for present purposes we may take to be verbally expressed, so that one is able to refer to or describe one's irritating aural sensations, with all that this implies about understanding. If I utter the words 'I have tinnitus', they are taken to be authoritative if given their conventional meaning, which is to say that my understanding of them is correct. Since understanding is partly dispositional, the implication here is that authoritativeness is compatible with dispositionality. It follows that if the authoritativeness of self-ascriptions of belief, intention and so on is called into question, then it would be inconsistent for this to be by virtue of the fact that their truth conditions have a dispositional component, for although experiential states are not themselves dispositional, the self-ascription of them *is*, for it involves understanding. It is a mistake, I believe, to assume that knowledge can be more certain than the understanding on which it depends and from which it derives its content.

This leads to the third point, which is that the distinction is superficial not only with regard to dispositions but in another way. If, again, I have ringing in my ears, my awareness of which is verbally expressed, then strictly speaking the reference is to a past event, for there is always a time lapse before the end of the sentence in which I refer to the sensations I am having. In any case, I cannot be aware of it now occurring without having the concept of it as a past event, otherwise I could refer to it only in the present tense and only until the ringing stopped. But now, if I imply knowledge of the past in ascribing a present experiential state to myself, and knowledge of the future, or conditional future, in ascribing forms of intentionality, so that both go beyond the present moment, then on that account there is no difference in the authority with which I self-ascribe them.

Observation versus introspection

If this is correct, then we still have a question as to whether *all* avowals, whichever category they fall into, are problematic in their authoritativeness. In separate chapters of *Knowing Our Own Minds* (1998) Wright and McDowell dispute some of the issues played out in this arena, and they disagree as to the paradoxical nature of authoritative avowals. Asking how it can be possible for avowals in either category to be non-inferential and authoritative, Wright detects a difficulty here which he takes the Cartesians to attempt to resolve when they assimilate introspection to privileged observation of one's own experiential or intentional states. This is to imply, presumably, that there are physical world observation statements that are epistemically secure in a way that the Cartesians take avowals to be. If they fail to make good the comparison, then perhaps this will be taken as confirmation that there is indeed a difficulty about avowals being authoritative. Wright's conclusion, in any case, is that 'a deconstruction of the privileged observation solution to the problem of self-knowledge is the indispensable prerequisite for an overall satisfactory philosophy of mind'. (p. 25) He takes Wittgenstein to have carried out that deconstruction.

McDowell disagrees with much of this, and he does not think that Wright has shown that there is a difficulty. Since I agree that he has not, and would argue that instead of a difficulty there are interesting questions to be asked, this being my position in the earlier discussion of Wittgenstein, I shall now take McDowell's approach a step further and try to show that observation reports cannot be used as a touchstone for the authoritativeness of avowals, and that these latter do not need to be compared to them. To begin with, it is clear that descriptions of observed physical objects, such as chairs, are inferential and go beyond what is immediately given. This raises a question as to whether non-inferential observations are possible at all, to which Wright seems to answer that they are, and he gives the example of a viewer looking into a kaleidoscope, which he thinks that Cartesians would regard as an analogue of introspective observation of one's own experiential states, at least if the viewer described only the pattern of colours and shapes. Given that the viewer is 'perceptually competent', says Wright, and is paying attention, then his description of what he sees will be authoritative. This has an odd ring to it – until, that is, one catches a glimpse of Wittgenstein in the bell tower, for Wright is presumably implying that there must be public criteria of perceptual correctness, otherwise it would not matter if, for instance, the viewer suffered from blurred vision and reported blurred patterns. That is, this would not matter if Wittgenstein could be ignored, perhaps silenced by the bells, for then it would seem natural

to us to say that others would accept the viewer's report, other things being equal, and that his own confidence in its accuracy would be the same as if he was clear-sighted.

Several points emerge from these considerations. If the viewer is concerned to tailor his description to immediately given colours and shapes, perhaps with visual qualia in mind, then his report is direct and non-inferential, just as it would be if he referred to his pain. The contrast here is with an account that quite clearly goes beyond what is given, as would be the case if he reported that he was looking into a kaleidoscope at objects reflected in mirrors at angles to one another, the patterns changing by rotation of the object chamber. But now, we should not read too much into this distinction, for it does not correspond to any fundamental epistemic difference. If he reports colours and shapes, or if a wasp crawls out of the tube and stings him on the eyelid, then both the visual and the pain report go beyond what is immediately given, and in ways that *are* fundamental. Thus it is that if Cartesians appeal to the notion of an observation report in the hope of raising the epistemic status of avowals, then they cannot succeed, and need not, for epistemic parity already obtains. Moreover, I have already argued that there is no significant epistemic difference between intentional and experiential avowals. If this is correct, and if it is nevertheless maintained, as in the passage quoted from Wright, that intentional avowals are suspect because they cannot be identified with self-ascription of experiential phenomena, then the rejoinder should be that the only significant difference this makes is that such avowals are not truth conditional upon phenomena of that kind.

But surely, it will be said, even if self-ascription of pain involves dispositions, the sensation itself is immediately and insistently there, in all its experiential urgency, which no expectation or other intentional state could ever be. Having prepared the ground for this objection, all I need do is to point out, as before, that the use of tenses is a necessary condition of declarative language. I grasp the concept of pain only if through change of tense I am able to keep track of it as being temporally positioned and having duration. If I can say that I am in pain, then I must be able to say that I was or will be in pain. This is necessary partly because I cannot capture immediate experience in my use of words. By the time I say that I am in pain, however fast my delivery, the experiential event to which I refer, referred, has come and gone, just as surely as if the reference was to a momentary twinge. This is to say that reference, even when it is in the present tense, is *transtemporal*.

Perhaps it is time to take stock. A cluster of issues attaches to intentionality and to avowals of intentional and experiential states, including a question as to the epistemic difference between the two categories. Having just addressed this issue, I suggest that there is

no special difficulty about epistemic access to our own beliefs, expectations, intentions and so on. That this should be in dispute is tied up with questions of privacy and semantic scepticism, which we have treated at length. There are puzzles about intentionality which I think can be resolved, now that it has been shown that they are not such as to be legitimately exploited by the sceptic. Of particular interest is the question of the links between occurrent and dispositional in a non-reductionist context, the answer to which will soon be forthcoming now that so much of the preparatory work has been done, with more to follow. In the meantime, and taking us in the same direction, I suggest that we now consider those aspects of avowals that make them authoritative, if indeed they are, and what it is that their authoritativeness consists in.

Avowals and the particular case of belief

To that end, let us focus on a particular intentional state, that of belief, and on the authority with which it is self-ascribed. The first step is to emphasise that avowals of belief are authoritative only if understood, though this is not to deny that in any individual case it is possible to hold a belief that may turn out to have been incorrectly or misleadingly expressed. For instance, if I claim to believe that I can count up to 100, then it is possible that there is a discrepancy between what I mean and what others understand, as would be the case if what I meant was that I could count up to 100 in tens, not in units. Even in such instances as these, however, I have to understand what I say, which I may not do, so that a question arises as to my authoritativeness with regard to it. There is a difficulty here which a comparison between belief and understanding brings out. If I speak in a particular case of believing that I understand, then the meaning is clear and unambiguous; but if I say that I believe that I believe that, for instance, I can count to 100, then either this locutionary device is pleonastic, nothing being added to the simple statement of my belief, or it qualifies that statement as being less than certain, more clearly conveyed by my saying that I think, but am not sure, that I believe that I can count to 100. Such uncertainty, however, is of no interest to us, for our concern is with simple, unqualified statements of belief. With regard to these, the expression of belief as an object of belief still has a use, but only in the past or future tense, in which mode it may still be somewhat circumlocutory, as when I state that I believe that yesterday I believed that I could count to 100. Still, this may unequivocally be referred to as a belief claim, since I claim that yesterday I had a particular belief, where the claim and the belief utterance are distinct entities. Nothing but confusion could arise, however, from merging them by referring to the form of words, 'I believe that…' as itself a belief claim when it

refs to itself, for this is to imply that one either has or pretends to have the belief, whereas it is also possible that one does not understand the belief sentence. There is a pitfall here, and clearly we do not wish the question of authoritativeness to fall into it, for the authority with which we self-ascribe belief would then turn out to be spurious. A similar risk would be incurred if we spoke of my sincerely claiming to believe that I can count to 100. If this is to imply that I am sincere in making the claim at the time that I utter the words, then what immediately follows is that the claim is true, since to be sincere is to have a belief.

In order to negotiate this terrain we need to start with 'I believe that…' regarded as an utterance or a sentence, so that we may trace out the epistemic and semantic conditions of its use as a statement of belief, thereby allowing for the possibility that they are not satisfied. A belief claim is then the proposition that the sentence expresses the particular belief, and it has truth conditions attached to it, such that oneself and others are able to judge whether it is correct. Suppose, for instance, that I utter the words 'I believe that I can count to 100', and let us call it sentence A. Then a concomitant of my having this belief is that I should be ready to oblige if asked for the immediate successor to any natural number between 1 and 99 inclusive. Now imagine that I do oblige but with the wrong answer; then it could be that I have made a slip, easily rectified, or perhaps the conclusion to be drawn after further testing is that I do not know how to count to 100. So my belief turns out to be mistaken, and therefore lapses, but I am still taken to have held that belief, for I was ready to oblige with answers and I understood the questions. Clearly, the belief claim would also hold good, and would not lapse, if the answers I gave turned out to be correct.

Now suppose that when it comes to it I am not ready to oblige with an answer or to start counting if asked, though perhaps I originally was. This is unlikely but at least conceivable, depending on the length of the interval between the belief ascription and the question, and if it does happen then it could be that I understand what is required of me despite being unable to comply. Realising this, I now say that I did believe that I could count to 100 but no longer do. In this case, as before, the belief claim is correct but my belief is not. In both cases there are possible circumstances in which originally my belief was true, my ability to count to 100 having lapsed in the meantime. If, however, the belief claim is not correct and never was, despite my having uttered the appropriate words, and if I was not feigning the belief, then we are free to imagine all manner of inconsistencies. It could be, for instance, that I am asked for the next number after 81 and admit that I do not know it, yet continue to utter the words 'I believe that I can count to 100'. In that case it may be suspected

that I do not understand my own statement, so that the question of authoritativeness lapses, since a necessary condition would not have been fulfilled.

What does this rather overcrowded account tell us about whether self-ascription of belief is authoritative in the sense of commanding acceptance by others? If to speak with authority is to issue a logical guarantee, then statements of belief must be certain in some other way, for they depend on understanding, which is not guaranteed, and on the dispositions appropriate to the particular belief. That is one answer, and what makes it more palatable, at least in the case of my belief utterances being judged by other people, is that there are strong inductive grounds for taking them to express belief, based on the frequency with which they turn out to be genuine. There is, too, a more difficult and revealing question, concerning my own authority for my belief claims, and in answering it we shall again enlist the notion of a system.

To begin with, it is worth noting that if I utter sentence A without understanding it, for instance because I do not know the meaning of 'to count', then this does not preclude the statement normally conveyed by it being true, so that I believe that I can count to 100, even if I would have to speak not of counting, were it outside my vocabulary, but equivalently of knowing the first 100 natural numbers in consecutive order. Interesting as this is, our concern is not with coincidence but with the use of sentence A to express the particular belief, essential to which is that I understand it. If I do not understand it, the question of authoritativeness with regard to it does not arise; if I do, then possibly I do not accept it, which is to say that I do not believe that I can count to 100. This again is to imply belief, albeit a negative belief, so that again authoritativeness is not in question; and this is also the case if I feign the particular belief, for again I am not mistaken in what I think, or really think, that I believe.

In order to drive a wedge between belief and utterance at a point where authoritativeness is put to the test I might ask myself how I know, having uttered sentence A, that I thereby expressed the particular belief. If, however, we suppose that the sentence was uttered some time ago, then this is the wrong question, for it concerns the memory of a past event and its corroboration. If, on the other hand, it was a moment ago, then the question would seem to amount to asking, effectively in the present continuous tense, how I know whether I have the belief, not how I know whether I had it in uttering the sentence. Could I not utter it and immediately think, 'No, that can't be right – I'm innumerate.'? Well, it is possible, except that this would be to correct the belief, not the belief claim. There is a difficulty here, and it connects with the fact, as it seems to be, that in any intentional state, in

this case that of belief, vehicle and content are inseparably interlaced, just as McDowell said. Perhaps, then, we should cease to try to keep utterance and belief separate, this being appropriate from a third-person perspective, as we have seen, but not from that of the person whose belief claim it is.

Accordingly, let us consider a straightforward question as to how I know that I have the particular belief, and try to give sense to it. After all, my belief that I can count to 100 involves dispositions, for instance my being disposed to be confident in attempting the count, so that I should be able to ask myself how I know that I would have that confidence. We have seen, however, when analysing third-person criteria pertaining to my self-ascription of belief, that there are complications. If I say that I would be confident in attempting the count if asked, then strictly speaking I have said too much, for my present belief does not depend on my future confidence, this being a condition only of my future belief. Suppose, then, that I utter sentence A and immediately am challenged to make the count, which I start to do and then stop at '3', unable to proceed further. What others make of this will depend partly on my explanation of it, together with background information; but ordinarily the presumption will be that my belief claim is false, the belief being feigned, and that I am innumerate. For myself, however, the epistemic access will be very different, for I will know, other things being equal, whether or not my belief claim of a moment ago was genuine – perhaps after '3' my mind just went blank.

Clearly, dispositions enter into the concept of belief in such a way that if a belief has occurrent status, perhaps by being verbally expressed, then one has to be suitably disposed, but always with its being possible that the disposition will fail when put to the test, provided that particular conditions are met. This, it seems to me, is a significant development in this discussion, for it accounts for and removes what strikes many philosophers, not just the sceptic, as the paradoxical nature of belief and other intentional states, in the light of what seems to be a discontinuity between occurrent and dispositional. How can I believe *now* that I can count to 100 if the truth conditions of that belief extend into the dispositional future? And how can my avowal of that belief be authoritative? The answer, as I have tried to show, is that the concept of belief operates within a system. If my belief that I can count to 100 becomes occurrent with my utterance of sentence A, then I have the belief, vehicle and content having fused into the one entity, and I am disposed to be confident and am confident at the time, though a moment later I lose that confidence, together with the belief. If that is the sequence of events, then perhaps it occurs within a wider narrative in which I tell myself, my mind having gone blank after counting to '3', that I have always known how to count to

100 and cannot understand how I could suddenly forget. This is to be contrasted with my feigning the belief and expecting not to be able to count beyond '3', that expectation again being non-reducible, occurrent in the given circumstances, and subject to dispositional constraints. If inexplicably I cannot count beyond '3', then this is, of course, an extreme case, contrived for a particular purpose; namely, to show that the appearance of paradox is illusory, the occurrent and the dispositional meshing together within a system.

Fundamentals in the context of a system

Let us now focus on dispelling the apparent paradox of the authoritativeness of avowals. If I believe that I can count to100 and suddenly lose that belief, together with the ability, or if I feign the belief, or if I do not understand, or fully understand, sentence A, then all this is possible only as part of a flow of beliefs, including perceptual beliefs, giving continuity to one's stream of consciousness, some beliefs remaining stable relative to others, which break and re-form around them. If particular beliefs or belief claims are buoyed up in a stream of beliefs which give them content, then this itself confers a kind of authority on them, belief being inescapable. What I shall now try to show is that this analogy can be made philosophical and supported by detailed arguments.

To that end, suppose I utter sentence A, or, just for a change, sentence B: 'I believe there is a swan in that box'. I have argued that either one has a particular belief or one does not, this being understood in terms of the difficulty of doubting one's own belief in a philosophically interesting way. Nevertheless, I think that I am able to find a vantage point from which to ask whether there *could* be reasons to doubt my own belief claim, hoping thereby to elucidate the nature of its authoritativeness. If I utter sentence B, thereby expressing that particular belief, which continues to obtain immediately afterwards, this does not preclude my stepping back and pointing out, for philosophical purposes, that it is logically possible, the uttering of sentence B having come and gone, that at the time, albeit only a moment ago, I did *not* believe that the swan was in the box, even though it seems to me that I did, and even though at present I have that belief. But now, the fact is that if in a particular instance I query my own belief claim, such that I at least pay lip service to bare possibility as a basis for doubt, as in the example, then this involves other beliefs, including my belief as to the content of the claim, a consequence of which is that the basis for my doubt cannot be such as to threaten *all* my claims to belief, as bare possibility would, otherwise my reasoning would fail to be coherent. It follows that if I state a belief, following which I cannot coherently entertain a doubt on the basis of logical possibility, then the statement of that

belief is made in the present continuous tense, duration being intrinsic to belief. To give another example, an essential fact about all verbally expressed belief is that it depends on dispositions and in that sense goes beyond the present moment. Therefore, it cannot be on that basis that I entertain a doubt in a particular case, nor can it be the basis for any general scepticism about what it is to believe or about the justification of belief claims. The reason is that such scepticism itself involves belief, a more general consequence of which is that it would be incoherent or self-refuting on *any* basis. It follows that a doubt can be entertained only in particular cases – not about my present belief, for if I have it I cannot doubt it, but about my belief claim, perhaps recently made, except that my reasons would be restricted in scope, and could not be such as to invalidate the doubt by way of the beliefs it calls upon. This is all very significant, as is the conclusion to be drawn from the discussion as a whole: that in particular cases self-ascription of belief may not be authoritative in the sense of being immune from doubt – but in the context of a system the self-ascribing of belief is *inescapable* and in that sense authoritative. It is part of the structure of the system and there is no escape from it in any form of discourse, including philosophical reasoning and the argumentation of the sceptic.

Clearly, these findings are relevant to the induction problem, direct consideration of which will recommence in the next chapter. At present our concern is with the part they play in a theory of intentionality in terms of a system, the issues raised by the sceptic always being in the background, and my aim in what follows is to tighten some of the threads of the discussion thus far. To that end, let us ask whether knowledge of one's beliefs can be assimilated to other forms of authoritative awareness, for instance to knowing that one is in pain, which again is something we would lay claim to, just as we are certain as to our beliefs. I suggested earlier, when countering Wright's comments on McGinn, that we self-ascribe pain by means of 'I am in pain' only if the sentence is understood, and that to understand it is to be suitably disposed. Following on from that, I would say that it is possible to be in pain without knowing it, as with infants and animals, and also possible, in one's ignorance, to utter the words 'I am in pain', as an injured parrot would. Conceivably, one could utter those words and know that one is in pain without thereby giving voice to it. For instance, such awareness would not be verbally expressed by a suffering monolingual Frenchman uttering the words 'I am in pain' when learning English phonetics. To forge the semantic and cognitive link between pain and utterance, we need to take account of dispositions, and in much the same way as with belief claims.

It is true, of course, that a belief is different from a sensation, which we know directly if it is the immediate experiential object of self-ascription in the present tense, as with 'I am in pain'. This belongs to the class of what Ayer refers to as basic propositions, characterised by their truth conditions being identifiable with the immediately given. These contrast with wider-ranging descriptions associated with physical object concepts, such as that of a rabbit, which one may understand while frequently confusing rabbits and hares. Although it points to an epistemic difference, the distinction thus illustrated does not, or so I have maintained, have the epistemological significance which some philosophers attach to it, claiming as they do that basic propositions are incorrigible. What is true is that if I believe that I am in pain, as expressed by 'I am in pain', then I understand what I say, otherwise no sense could attach to my believing it; and then it follows that I am in pain, for I do not understand if, without pretending, I say that I am in pain when I am not. But now the question shifts from how I know that I am in pain to how I know that I understand. Here we should keep in mind the claim, which Wright advanced in the quoted passage, that my knowing that I am in pain is directly grounded in the sensation itself, an example, perhaps, of what Ayer refers to as an 'act of knowing' (1959, p.111), a form of words which he thinks is highly misleading. Indeed it is if it is taken to exclude dispositions, for reference to one's own pain is not self-contained, given that one's use of 'I am in pain' is correct if and only if one is in pain, so that on any particular occasion of use one has to be not only in pain but also suitably disposed with regard to one's use of 'pain'. But 'an act of knowing' connotes direct awareness, whereas nothing counts as being directly aware of a disposition, which is in a different category from sensations.

If the foregoing is correct, and if my use of 'I am in pain' is considered, then I may not understand it even if I think that I do, for that understanding depends on my being suitably disposed with regard to the use not only of 'pain' but also of the other words in the sentence, all against a background of general competence in English. It is not possible to specify exactly how far my command of the language must extend, partly because understanding admits of degrees, which is why I have concentrated on essential cognitive and semantic conditions. One of these is that I must be able to repeat or paraphrase the sentence, other things being equal, and what this indicates, as so much else does, is that the required dispositions are utterly dependent on my memories being correct, which again invites a question as to how I know that I understand.

Does this mean, if the pitfall is avoided of pre-supposing that one is in pain, that in fact we cannot know that we are? Not at all: in starting from 'I am in pain' regarded as an

utterance, all that we have done is to reject a particular view of what it is to know that one is in pain, that which presents awareness of pain as grounded in the sensation itself, no account being taken of dispositions; and a corrective to this, as we have seen, is to analyse that awareness in terms of a system. But that analysis enables us to reinstate authoritativeness on a different footing, very much in the same way as with self-ascription of belief. If we consider, for instance, the epistemology of memory, then it is true that self-ascription of pain is conditional upon one's memories being correct, which in particular cases they may not be; but what must also be said, by way of balance, is that it is in the nature of the system to impose severe constraints on any doubts we may have as to the legitimacy of memory claims. These are such as to rule out general philosophical scepticism about memory, given that the sceptic argues from within a system in which no statement can be made that does not pre-suppose that the memories it calls upon are correct. Reliance on memory, as was asserted earlier about self-ascription of belief, is part of the structure of the system; and this is to imply that if on a particular occasion I doubt whether I remember correctly with regard to my understanding of 'I am in pain', then my reasons must be appropriate to the particular case, not general reasons such as to render incoherent the doubt itself. It could be, for instance, that owing to cognitive deterioration I am no longer certain as to the meaning of 'pain', having lost confidence in particular linguistic memories – though not, let it be noted, in all of them, otherwise I should not be able to verbally express my misgivings about some of them. The point, anyway, is that my loss of confidence would debar me from stating that I was in pain, or that I knew that I was. It is conceivable that I confidently utter the words 'I am in pain', as if self-ascribing pain, when in fact I am not in pain; but then I do not understand what I say, so that I am not mistakenly self-ascribing pain.

Furthermore, if scepticism about the understanding of 'I am in pain' is based on the argument that it is logically possible that one utter that sentence without understanding it, the necessary dispositions bring absent, then the same is true of any sentence, including that by which the argument is conveyed. It follows that if I state that I am in pain, and then it occurs to me as a philosopher to ask how I know that I am, then this immediately rules out any sceptical answer based on logical possibility.[5] But further, and as already explained, if I can ask how I know that I am in pain, thereby exhibiting my grasp of the concept of pain, then it must be that I am in pain when I say, without feigning it, that I am, so that the question answers itself.

---

[5] *Pace* Descartes.

This should all be enough to establish the qualified authoritativeness of pain avowals, depending as they do on dispositions, especially if it is noted that a condition of one's grasp of the concept of pain is only that one should be suitably disposed at the time, not that one's dispositions should permanently obtain. Thus, it is possible that I know on a particular occasion that I am in pain, but that my understanding then lapses, so that I am no longer able to refer to my pain on other occasions, perhaps because I have forgotten the meaning of 'pain'. It is a significant fact that one has to make an effort to imagine such a thing, for it indicates how well entrenched is our grasp of basic psychological concepts, such as that of pain. There are inductive connections here, which await investigation; but for the moment, and by way of gaining insight into the entrenching of such concepts, we may note that the grasp of them is taken to extend over the continuous present. That is, one can imagine that my claim to know the meaning of 'pain', to know it now at time t, is tested by my being stimulated in various ways and asked to describe my sensations, and, less painfully, by my being asked what I infer from the screams of other volunteers. It may be, as in practice it almost certainly would be, that I am deemed to have passed the test. But this all takes time, so in a sense the verdict is retrospective; or, rather, it is taken to cover the period between time t and the continuing present. Now, it might be thought that this reveals only that one's abilities tend to persist; but I would argue, as in the next chapter, that persistence of this kind, with its inductive implications, is a fundamental semantic and cognitive condition and is part of the structure of the system of language use.

It will be instructive if we now compare belief and understanding, for just as belief is all-pervasive, so is understanding. If I believe that I have grasped the concept of pain, as exhibited in my use of it on this particular occasion, so that I am at present suitably disposed, then as with belief avowals there are various possible scenarios. It could be, for instance, that when I next use the word 'pain', and in accordance with my present disposition, it turns out that my use of it is incorrect, a mistake which perhaps another person draws to my attention when my feet are tickled and I say that I am in pain, and please don't stop. I now revise my use of the word to bring it into line with conventional usage, the point being that both my incorrect and my revised uses are possible only within a framework of understanding on my part. As with belief avowals, then, it is only in particular cases that doubt can intrude. I cannot entertain such doubt on the basis that a mistake is logically possible, or that understanding is not only occurrent but also dispositional and therefore holds the future to account, or that it depends on memory or in other ways goes beyond what is given – and the reason is that all this is also true of the understanding which I take for granted in entertaining

the doubt. To attempt to deny that one understands one's own sceptical settlements is to lapse into incoherence.

What I now propose is to weave together some of the strands of the discussion thus far. In an earlier section I tried to correct what seem to me to be misconceptions about the difference between experiential and intentional states; for instance, I pointed out that although the present tense self-ascription of pain refers to a present sensation, the concept of pain requires that one refer to it in the past and future tense, when it is not experienced, and that one refer to other people's pain. It is worth remembering, too, that one can feel pain without being able to refer to it, as with children and animals. To make these points is to minimise the role of immediate sensation and in that way to assimilate the experiential to the intentional. To that end, I also tried to show that it is evidence of ontological bias if intentional states are regarded as lacking in conscious content so that they are somehow less real than sensations and perceptual experiences. These points having been made, I am now able to say that intentional states are not abstract or ghostly or in need of reduction to ontologically respectable experiential entities. It was partly this bias against the intentional that Kripke was able to exploit in his semantic scepticism, and in his dismissal of the possibility of intentional states being non-reductive.

The thesis that the conscious content of intentional states has ontological parity with that of sensations in respect of being captured by concepts and entering into discourse is one that intersects with everyday experience at many points. If, for instance, a belief is made occurrent by being verbally expressed, then the conscious quality of the belief utterance is different from one's consciousness when the belief is feigned, itself distinct from one's awareness of not understanding the utterance, as with a French student learning English phonetics, this itself being different from thinking that one understands but not being sure, as when struggling to work out a formula or grasp a philosophical problem. Then again, there is what it is like to recognise an object as a table and move smoothly on to the next article of furniture as one's eyes sweep a room; and now contrast this with feeling uncertain about whether it is a table, or with the sudden realisation that it is a worktop fixed along one side to a wall, all of which may be verbally described. And, too, a familiar experience is that of descending a flight of stairs on automatic, as it were, only to be forcibly confronted with hard reality when one steps down onto the last but non-existent tread and struggles to regain one's balance. Finally, we must not forget propositional attitudes, as when I think that people in the next room are talking about me. There are interesting differences between these disparate examples, but what they all indicate is that belief, expectation and understanding cannot be

separated out from the continuous flow of experience, in which the eddies of surprise and uncertainty are caught up, as we are swept along from one moment to the next. Rather, they partake of that experience, or they inform it, the expression 'raw experience' being an oxymoron, and the associated dispositions find their actualisation in the same flow of consciousness and awareness.

Irreducibility again

Perhaps it is time to more explicitly incorporate irreducibility into the continuing development of this analysis of intentionality, and in particular of meaning and understanding. According to McGinn, these latter are primitive, non-experiential states; but it seems to me, in line with the previous arguments, that what it is to understand is not so easily separated from, for instance, the experience of flicking a switch and expecting a light, or seeing a table as such, or listening to a piece of music, or looking at a sequence and trying to work out an nth term formula, or reading a text, or registering one event as immediately following another – all of which, in my view, are informed by understanding. As for primitiveness, there may be nothing to object to here, but I prefer to identify irreducibility more closely with the arguments in support of it. If those arguments are correct, then meaning and understanding are irreducible in the sense that they cannot be captured by any definitional account that does not employ semantic and cognitive concepts. This rules out any such account in terms of dispositions, the place of which is in analysing, not reducing, the conditions attached to what it is to mean and to understand. Thus, the understanding of the even number sequence includes being disposed to expand it, a process which is informed by understanding, itself dependent on dispositions, the actualisation of which is again informed by understanding. There are connections, too, with memory and the concept of the past, that concept itself being irreducible and capable of application only if informed by understanding. In other words, nothing counts as dismantling the relevant dispositions and laying bare the non-cognitive elements of which the understanding of the even number sequence is composed.

Perhaps further insight may be gained if we draw a parallel with perceptual awareness, itself a form of understanding. Suppose I say that I see this table, so that my looking at it is a present occurrence; then my identifying it as a table goes far beyond my perceptual experience of it at a particular moment. I imply, for instance, that its visual appearance changes with the position of an observer, in this case myself, so that I am disposed to expect those changes. But now, it is intrinsic to my awareness of the table that I

'see' it as having three-dimensional depth, an effect I can eliminate only by screwing up my eyes and losing sight of the table, or at least of the table as such. It follows that I could not do justice to my conception of it by describing it in terms of two-dimensional coloured shapes. If this is correct, then just as we 'grasp' the essence of a rule in applying it, so it is that we 'see' the three-dimensionality of a physical object visually perceived, or 'feel' it at the tactual level, the inverted commas in both cases marking what appears to be problematic and yet is essential to the concepts. My experience of 'seeing' the table is immediate and direct, and it obtains within a dispositional framework. A disposition is not itself an experiential object, and yet, it is as if my 'seeing' is intrinsically revealing of spatial depth, hence the comparison with grasping a rule. To speak in this way is to seek to express an elusive truth; but I would say that it does not incur the risk, as Wittgenstein would have it, of being misled, and that there is no point in galloping to the rescue if no-one is in difficulties. We would all agree, for instance, that 'seeing', insofar as it reaches beyond itself, can guarantee only inductively, not mystically, the appearances that experience has taught us to expect in any particular case. Similarly, and in any individual case, to think that one understands is not to understand. Our concern should be with analysis, not the dispelling of phantom misconceptions.

To that end it should be noted that the appearances in question, for instance of a table, are themselves characterised by spatial depth, so that what I expect is continuity of 'seeing' the table through changes of colour appearance and shape. It follows that if the experience of 'seeing' three dimensions informs my changing visual perceptions of a physical object and its surroundings, then my belief as to what I see obtains within a system in which the existence of the physical world is given. It is still true, of course, that in any particular case the correctness of such beliefs is not guaranteed, there being only a contingent link between the physical aspects of an object which change of perspective reveals. But this is within a system in which belief about the changing appearance of an object viewed from different positions, or about changing physical events or processes from a static viewpoint, is connected with other possible beliefs as truth-conditions of having that belief, where all such possibilities involve the 'seeing' of spatial depth. Thus, if I believe this object to be a table, viewed from the front with two legs showing (the table's, not mine), then with relative change of position I would expect to 'see' two back legs; but my being so disposed is a truth-condition of my believing that this is a table only if on the basis of 'seeing' those legs I believe that they exist, which itself implies other beliefs occasioned by further change of perspective relative to the legs.

It follows that a condition of my believing that this is a table is that I should be disposed to have other beliefs pertaining to it, such that my perceptual awareness of the table obtains within a system of interconnected beliefs, none of them reducible, given their interconnectedness, and each of them based on observation, so that they are both occurrent and dispositional. This encapsulates one of the main points needing to be emphasised, and implicit in much that has already been said; namely, that when I use an expression my dispositions determine not only *whether* I mean anything by it but *what* I mean. Thus it is that the particular meaning of 'this is a table' is not reducible to dispositions, which themselves involve meaning and understanding, but is determined by them, the dispositions themselves being determined by my cognitive and perceptual interaction with the world. Just as a particular belief is not reducible to the attaching dispositions, which are actualised by further beliefs, by those implied by the original belief, so it is that to understand is to be disposed to understand, the system thus constituted being one in which the semantic and the cognitive are irreducible. It is only in this sense, not Wittgenstein's, that meaning and understanding may be resolved into practice.

Let us now enquire more closely into the ways in which dispositions are determined. If I say 'this is a table' and mean it, then I expect when suitably positioned to see two back legs holding the table up; otherwise, it is a worktop or counter, or has gained exemption from the laws of gravity. This connects the semantic and cognitive with the epistemic, and it illustrates the general point that to know what an expression means is to have knowledge of the object to which it refers. With a table one need only recognise it as a four-legged, flat-topped article of furniture, but a more advanced intellect is required in order to recognise the objects, namely people, who may be seated around it. If one of them is reading aloud from a newspaper, which I am reading over his shoulder, then I would explain the match between his reading and mine by inference to his consciously scanning the text, just as I myself am doing. *Pace* Wittgenstein, this explanation enters into my understanding of 'he is reading aloud', and this is an example of the interplay between linguistic and empirical understanding, which both have a role in this application of the concept of other minds. Thus it is that the cognitive, semantic and epistemic permeate one another within an explanatory system in which dispositions prominently figure.

Since this is the penultimate chapter of the induction part of the thesis, I shall not attempt to recapitulate its main points, as these will be included in the general summing up at the beginning of the final chapter. What I propose instead is to venture a conclusion, one which derives from placing the epistemological sceptic in the world as a user of language,

without which there could be no scepticism. He is forced, thus situated, to desist from any attack which would rebound against himself, so that with most of his arguments neutralised he is unable to challenge the legitimacy of belief claims or knowledge of the past or of one's own language-related dispositions. But this, surely, is to give shape to a resolution of Hume's problem itself, for we know that the inductive sceptic is obliged to issue such challenges. What remains is, as it were, to completely unveil the solution, the main outlines of which are clearly visible in the folds of the discussion thus far; and this is the task to which I now turn.

Chapter 6

Recapitulation and Solution

My aim in this final part of the thesis is to mark particular strands from each chapter with the same dye, that of essential relevance to the induction problem, and then to weave them into a refutation of inductive scepticism. The first chapter was devoted to an examination of Keynes' analysis of probability in terms of a relation of partial entailment, one by which the equipossible becomes the equiprobable, the calculus of chances applying to the alternatives thus derived. That analysis, if it could be carried through, would supply logical grounds for inductive inference. I argue, however, that the notion of a quasi-deductive probability relation is too problematic, and that in any case the inferences into which probability relations enter have premisses in which the validity of induction is presupposed. The significance of that fact is, of course, a matter for investigation.

As a first step, I now consider the adjacent question, as it were, of the logical status of so-called synthetic *a priori* statements, such as those of applied geometry, and I then discuss the simpler case of linear inequalities, agreeing with Ayer that if logical truths are factually interpreted then nothing is allowed to count against them. Instead of leaving it at that, however, I now point out that applied logic – or mathematics – is possible only within a system in which the concept of numerical or qualitative identity is entrenched. This is also a condition of inductive inference, which relies on repeated instances, and of what it is to refer, since the reference must be repeatable, and words have meaning only if re-usable, with all that this implies about recognising them as the same. But now, these conditions govern pure as well as applied logic, both of which, in addition, depend on memory and the concept of the past, as does inductive inference. There would therefore seem to be a level at which essential aspects of induction are fundamental, without which it would not be possible to reason either inductively or deductively. If this is correct, then we need to ask whether it points to a solution to Hume's problem, a question I do not try to answer in this first chapter; what I suggest, instead, is that we keep it in mind when considering other attempts at deploying probability theory against inductive scepticism.

This brings us to the second chapter, on inverse probability and Fisher's anti-sceptical views. He claims that logical relations enter by way of intrinsic probability into the no-miracles argument, which thereby provides grounds for inductive inference; and he tries to demonstrate this by considering cases in which a hypothesis leads to successful predictions. Again, then, we have a possible rejoinder to Hume, except that I now take issue with the very idea of intrinsic probability. My approach is to differentiate it from the appeal

to direct evidence, which occurs when, for instance, we derive a probability measure from the colour proportions of balls in a bag. With regard to the colours to be expected when drawing from the bag, such measures have predictive value only within a system in which they are bound up with frequency. This includes the frequency of correct guesses which validates the no-miracles argument, if a systematic approach is taken to its application to cases of inverse probability in the absence of initial probability values.

The appeal, then, is to direct evidence, but only in the context of a system; and this is the location, also, for the related fact that probability is relative to evidence, this being the kernel of truth in the appeal to logic of both Fisher and Keynes. It is *because* the balls in a bag have a particular colour proportion that I assign a corresponding value to the probability of a given colour being drawn. But this is within a system in which probability is linked to frequency, my awareness of which may or may not be part of my understanding of the concepts involved. That system is characterised by interconnectedness: frequency data require interpretation by statistical methods involving inverse probability and the no-miracles augment, which themselves are validated in terms of statistical frequency.

And so to chapter 3, which begins with a critique of Stove's attempts to use formal probability arguments against Hume. I try to show both that Hume's scepticism cannot be fitted into a formal probability system and that Stove's anti-sceptical arguments are in any case invalid. Having disposed of attempted probability solutions, I now focus on what I think is the correct approach, for which some preparation has already been made, and I enquire first of all into the epistemic status of numerical inequality in probability and induction. If a bag containing red balls and black balls has a higher proportion of the former, then it is rational to expect to draw a red; but this is the case, or so I claim, only if that proportion correlates with relative frequency in a series of trials. This is to say that I treat frequency and proportion differently, requiring frequency to evidentially validate proportion, despite their being numerically the same, a fact which the sceptic may turn to his advantage. He may challenge us to explain why numerical inequality should be regarded as being epistemically significant, thereby assuming the role of precursor to Hume; and in my case he may ask why it is that I attach a significance to frequency which I withhold from proportion.

What I now suggest, if answers are to be forthcoming, is that we consider the fact, already alluded to, that probability and induction operate within an epistemic system in which the link between evidence and frequency is pre-supposed and enters into the premisses of any factual inference. And in particular, or so I claim, it enters into physical world description and belief about the past, a question now arising as to whether there are *any* areas of factual

knowledge that may prove to be independent of induction once we delve into them, the answer to which is perhaps that it depends on the depth of excavation. Rather than going deeper, at least at this point, I now make a pivotal and complementary claim: that the narrow Humean challenge to induction, to those beliefs that are justified and generated by appeal to past instances, is inherently unstable and could achieve stability, if at all, only by occupying a wider base.

In support of this claim, I take each of Hume's arguments against inductive inference and attempt to show that if they are valid they also apply elsewhere, thereby providing further reasons, not circumscribed by the focus on induction, why physical description and belief about the past must also be targeted by the sceptic, those reasons being such as to bring experiential statements, too, within range, so that all areas of factual knowledge should be taken to be under threat. Experiential statements are at risk because semantically and cognitively they depend on dispositions, and in that sense they go beyond the immediately given, so that it is a mistake to believe that they command an impregnable epistemological position. That, at least, is the point of view I propose for consideration, and I continue to argue in support of it in subsequent chapters.

If my claim about the reach of the inductive sceptic is correct, then he can be coherent, if at all, only by yielding to global scepticism, into the arms of which his arguments propel him. What now needs to be shown is that those arguments cannot live and breathe in such a tight embrace, for he is unable to formulate them without accepting restrictions on what it is possible to coherently doubt. The reason is that words have meaning and are understood only if the existence of the past and constraints upon the future are pre-supposed. That, at least, is what I now intend to establish, in chapter 4, by immersing myself in the intricacies of intentionality theory.

In that chapter, then, I start with Wittgenstein and take issue with his thesis that inner processes stand in need of outward criteria, which I suggest is no more true than its converse, so that I set it against the notion of an explanatory system connecting the inner life and outward behaviour of myself and others from a first-person perspective. On that view, intentionality involves both inner and outer, as when my understanding of the even number sequence, for instance, consists in my ability, among other things, to supply the successor to any term, which itself involves my recognition of even numbers as such, itself an inner process, and my being suitably disposed in a variety of ways, with inner processes again involved when my dispositions are made actual.

Our aim, it seems to me, should be to delve deeper into all this, so that we analyse intentionality by ascertaining its essential components and tracing out their interconnections. It is in the nature of the case that the different factors must be compatible with one another, and this constrains the account that we may give of them. Thus, in this chapter I focus on the occurrent and the dispositional, and in the next I include irreducibility, the fact of their co-existence helping to determine what is to be meant and understood by the links between them, consideration of which leads to my use of the notion of a system.

Set against this, however, are Wittgenstein's views on the confusion surrounding the nature of intentionality, as personified in the form of the interlocutor, who is aware that understanding has a dispositional element but clings to a picture of it as being self-contained, so that writing down the first terms of the even number sequence is evidence of understanding in the same way that yawning may be evidence of tiredness. In his befuddlement, having slept badly, he would also incline to the view, perhaps, that understanding generates its own application, though without being able to point to a causal inner state, the counterpart to feeling tired.

That, at least, is an interpretation of Wittgenstein's views which makes sense of them, except that it raises a question as to their import. It is obvious, after all, that application is integral to meaning and understanding, which therefore cannot be self-contained. One could think of them as inner states which are manifested in application, so that there are points of analogy with the ways in which the occurrent and dispositional enter into such concepts as, for instance, that of brittleness. Given, however, that we are likely to be aware, also, of the points of disanalogy, it is not at all clear that we need to be delivered from any misconceptions arising from the comparison. As for the parallel drawn with tiredness, we are surely not tempted by the assimilation of understanding to sensation or feeling: what, after all, could be more different?

A query must be raised, then, as to Wittgenstein's professed aim, which is to lead us out of confusion, some doubt arising as to where exactly he is taking us. He claims to leave everything as it is, but his thesis equating meaning with use has been radically interpreted. Anything more moderate, or so I argue in the next chapter, and that thesis reduces to platitude. What I have tried to show in this one is that the equating of meaning with use, if interestingly understood, follows not at all from Wittgenstein's arguments as to the ways in which we are misled. Hence the detailed discussion of his views.

Finally, then, we have the chapter before this one. I begin with a rebuttal of meaning as use, in association with which I take the realist view, with regard to asserting that a past

event has occurred, that what makes the assertion true is just the occurrence of the event, so that the concept of the past is irreducible. On this view, such statements are connected with use only in the trivial sense that they are used to refer to past events. That, at least, is what I maintain, and it conflicts with Kripke's radical scepticism, the essence of which is that meaning and understanding are non-factual, so that sameness of meaning has to be construed in terms of the practice of a community. For Kripke, the appeal to such practice furnishes what he refers to as a sceptical solution, the inadequacy and incoherence of which, in my opinion, is easily exposed. For what it implies is that nothing counts as violating a meaning-rule, provided that any change in practice is common to the community. Clearly, there is a conflict here with what is essential to the notion of rule-governed action or meaningful discourse, hence the incoherence.

Having repudiated the sceptical solution, I now argue that the sceptical conclusion it attempts to circumvent must also be fatally flawed; and again this is easily shown. If the claim is that words have no meaning, then how are we to understand the semantic status of the claim itself? We cannot say that it has no meaning, or that there is no fact of the matter as to what it means; therefore, it is self-refuting or incoherent.

What we are left with, now that the semantic sceptic has been shown to be irredeemably confused, are those features of language use found problematic in both Kripke and Wittgenstein. By way of illustration, if I point to a blue object and say 'this is blue', and if a question arises as to what constitutes my meaning that this is blue, then a phenomenological description of my conscious state will not suffice as an answer. For instance, my visual experience of the blue object is not enough, not even when conjoined with my saying 'this is blue', for I may be a foreign student parroting the words. Time, then, to call upon dispositions, except that Kripke would object that semantic concepts are normative, so that my actual dispositions may not be the ones that I should have if I am to mean that this is blue. To this it may be added, by way of reinforcement, that I would claim to understand my words *now*, and yet a disposition is purely abstract and conditional.

We seem to be at an impasse, and the way out of it, or so I now claim, is to follow McGinn in appealing to the notion of semantic and cognitive irreducibility. This, in my view, transforms the problem by re-directing the enquiry, the focus shifting from what it is to mean and to understand to the ways in which they operate within a system in which they are occurrent, dispositional and irreducible, these basic properties and their interconnections now needing to be explored instead of challenged.

This difference in approach makes itself felt when I now turn to the analysis of avowals in rejoinder to Wright's criticisms of McGinn's irreducibility thesis. I argue that such factors as the dispositionality of self-ascription of belief do not render it less authoritative than is the case with basic propositions, though we need to be clear as to what it is that such authoritativeness consists in. I proceed to analyse such ascriptions in terms of their place within a system, giving due emphasis to the distinction between belief utterance and belief claim; and I point out that I court inconsistency if, as a philosopher, I entertain a sceptical doubt as to whether my beliefs are what I say that they are, as opposed to my not being sure, in the everyday sense, what it is that I believe. The reason is that the querying of one belief ascription implies the acceptance of others, so that the grounds for doubt cannot be such as to hold universally, as they would do if appeal was made to such facts as that belief is dispositional, or that it relies on memory, and thereby goes beyond the immediately given.

Solving the problem

Since I have now completed this short survey of the more prominent peaks of the discussion thus far, all that remains is to travel the route which connects them and leads to a solution. Our starting point is the Humean view of inductive scepticism as, essentially, rejecting all knowledge of future events or unobserved present events, so that the threat posed by the sceptic, however alarming, is taken to be restricted in scope. It may be argued, I think, that if we concur in this as being a coherent view, then the sceptic will always triumph over common sense. To see why this should be the case, we need only recur to the earlier discussion of the validity of probability judgements that are disengaged from frequency. I claimed that if all that is known about a single ball in a bag is that it marked with a number between 1 and 100, then nothing follows as to whether not-1 is more probable than 1. This changes, or seems to, when proportion is linked to frequency in a series of trials, as when I am allowed to choose which number to bet against, for in a sufficiently long series of trials I can expect to win about 99% of the time. There are two considerations here, the first of which, given that the Humean sceptic rejects probability as a logical relation, is that he would deny that such expectations can be justified. What must also be considered, and is often overlooked, is that even if such expectations are given a temporary licence, he may still maintain that I have no grounds for expecting to win on any particular occasion, and for the reason that I should have to rely on the epistemic value of numerical inequality[1]. He may

---

[1] See the discussion of Okasha, about to begin, and the definition of narrow inductive inference.

point out, too, that this argument extends much further than one might assume, applying as it does to cases in which *all* instances are favourable, as with expecting to draw a black ball from a bag containing only that colour. After all, even if the balls were black when placed in the bag, there is no reason to think that they still are, for any such reason would be inductive in character.

This last example should give us pause, but if nevertheless the Humean sceptic's terms are accepted, so that his demolition of our beliefs is taken to be limited in destructive force, then we may come to see with what seems to be great clarity, once the dust has settled, that we have no reason for any course of action, or to expect one thing rather than another, apart from what Hume vouchsafes to us in the form of habit and custom – except that they, too, are unsupported by reason. It is important to yield to this negative view, albeit momentarily, and to realise that a void of pure nescience opens up before us where the future should be, forcing us to retreat into the past though corridors of memory left open by the sceptic. Thus, we are permitted to say that the sun has risen in the past but not that it will rise tomorrow, or that black balls and red ones were placed in a bag, with red predominating, but not that we may expect that colour to be drawn. This is a bleak view of what it is possible to know, and in my view its hold on us may be broken only if we challenge the sceptic's self-imposed limitations and show that the imperatives of Humean scepticism lead to a wholesale rejection of factual knowledge.

That a resolution of Hume's problem depends on widening its reach is a proposition that some philosophers would dispute, arguing as they do that his attack on induction may be met head on. I have criticised such confrontations with the sceptic when they depend on probability theory, but I would now like to consider the oblique approach adopted by Samir Okasha (2001), whose thesis is that Hume proved only that inference from the observed to the unobserved is not deductive and therefore is not infallible, which we already knew it not to be. According to Hume, he says, such inferences are not just fallible but entirely without foundation, for if they are not demonstrative then they must depend on past experience via a principle of uniformity by which regularities are projected; but that principle could be justified only by appeal to past regularities, which would be question-begging. The flaw in that argument, claims Okasha, is that no such principle is in use, for we project some regularities but not others, the link with expectation and prediction always being forged so as to agree with background beliefs. He does not attempt to explicate Hume's argument, which he acknowledges has been variously interpreted, but he would agree, I think, that we speak of a principle of induction or uniformity when we are concerned with rationalising the inference

from past regularities to unexamined cases. That is, we conceive of such inference as being linear, epistemically separate from any background information, and based solely on repeated instances which determine the nature and strength of the resultant belief. Thus characterised, it is easy to see how it could be made to fit into a formal probability or inductive logic system. Life, however, is messier than maths, but this did not stop Hume, or so it seems, from assuming that we reason inductively in this narrow sense. The point, anyway, is that to speak of a principle is to float the idea that such inference is in some sense valid, but it is not in itself to add any support to that idea, so that we beg the question, just as Hume says, if we appeal to that principle in defence of the idea.

There is no reason why Okasha should disagree with this, and he now suggests, as already mentioned, that if actual empirical reasoning is not inductive in this narrow or Humean sense, then perhaps Hume's sceptical argument is valid but not sound, for it is premissed on an incorrect account of the inferential link between past observations and present expectation. He favours a Bayesian account of belief formation such that we start from subjective prior probabilities conforming to the probability calculus, and then we learn from experience by a process of conditionalisation with respect to empirical data. To criticise Hume in this way is to run counter to the views even of some of those philosophers who would agree that narrow inductive inference is rarely or never met with, for they could still insist that the sceptical argument is effective against any form of factual, which is to say non-deductive, inference, including Bayesian probability reasoning. This is the position taken by Mark Lange (2002) in his reply to Okasha, and no doubt he shares it with others; so perhaps we should join the debate and enquire into the soundness of Hume's argument, assuming this time that induction in the narrow sense just defined does not in practice contribute to factual belief.

In chapter three I made out the case that inductive scepticism entails global scepticism, my arguments being predicated not at all on the accuracy of Hume's account of induction, which I attempted to undermine by showing that past regularities are themselves vulnerable to sceptical attack. In asserting that the scope of the Humean sceptical thesis is unrestricted, I took its essence to be as follows: factual inference is non-demonstrative; also, any attempt to justify it in general terms would presuppose its legitimacy; therefore, it cannot be justified at all. Strictly speaking, I suppose, one could simply maintain that factual inference is legitimate and leave it at that, just as Moore held up his hands in order to vindicate the common sense view of perception to which we all subscribe. This semi-naked appeal to what we all know to be true outside the study does not meet with much approval within, or not unless it is able to slip into something more suitably philosophical, a similar

response to inductive scepticism also being deemed to be inappropriately dressed. That being the case, I did not consider such a response but instead went on to extend the range of Hume's thesis, for instance by arguing that knowledge of the past is contingent and depends on memory, as does any attempt at justifying that knowledge, which Hume is therefore obliged to reject. The point I am trying to make, then, is that the soundness of Hume's argument does not depend on the adequacy of his account of factual reasoning; and I am about to try to show, of course, that it *does* depend on not entailing global scepticism, a condition it fails to meet. Speaking of which, both Okasha and Lange presuppose in their discussion that Hume's sceptical thesis does not, as it were, go all the way down to the ocean floor, where no factual belief can be kept afloat and all empirical reasoning has drained away. They take for granted that basic empirical information is not itself in peril on the sea, but only the derived factual beliefs concerning events over the horizon, and the waves of inference propagated in between; hence Okasha's appeal to Bayesian methods of belief acquisition. But those methods cannot find any firm anchorage – or not until the tempest of Humean scepticism has been tamed.

In manipulating the epistemological weather, waiting out the storm not being an option, I shall take for granted, the case having already been made, that Hume's arguments are global in their reach, so that they threaten not only knowledge of the past as well as the future but also what we know of the present, in the form of observation and experiential statements. That given, and if we are to demonstrate that global scepticism is self-refuting, then the place to start, which in a sense is where the induction problem originated, is in Hume's study. We may imagine him ensconced in his reflections, quill in hand as he ponders a paradox of inductive scepticism: that there can be no reason to perform any action, there being no reason to expect any particular result; and yet, it is impossible not to act, and always in one way rather than another, one's actions arising from and manifesting one's beliefs, the grounds for which collapse under philosophical scrutiny. Hume's sceptical resolution of the paradox, reminiscent of Kripke's in a different vein, is taken to be that he appeals to habit and custom, except that he makes no express attempt to justify induction by making it feel more comfortable, as it were, in the familiar surroundings of one's past history of inductive belief. This, from our perspective, is just as well, for Hume can be consistent only by rejecting that knowledge of the past which is presupposed in any appeal to habit and custom. Setting that aside, it is clear in any case that any such appeal would fudge the issue, the firepower of his sceptical arguments being directed at demolishing any appeal to past regularities not only of events but also of inference. From those arguments it follows that the accumulation of past

instances of an event gives no reason to expect future ones; and, concomitantly, that no reason is given, either, by past instances of *expecting* an event, not even if the expectations have always been confirmed, as with the rising of the sun, such confirmation being just as epistemically barren as would be the case if one's predictions had invariably failed.

A rejoinder here, perhaps, is that Hume is not attempting to justify inductive practice, not even by the back door, but simply wishes to draw attention to it as being natural and inevitable. This, it seems to me, is disingenuous, for subliminally the appeal to what is natural takes nature to know best, so that a picture flashes onto one's mental screen of our having evolved a form of reasoning which engages with a world in which past regularities are reliable indicators of future events. But this, as I said, is to fudge the issue, for Hume cannot expect such a world to continue. The fact is that if we have no reason for any action, or for the associated beliefs, then we have *no reason.* To deny this, in other words to equate reasoning with habit formation, is to reduce the rational to the mechanistic, and humans to the level of carpet mites– there being, however, a fatal flaw in that equation, for arachnids do not philosophise, and then there could be no problem of induction.

What is also clear, I think, is that many passages in Hume, even if he is taken to be a restricted sceptic, not a global one, are problematic in another way. He claims, for instance, that when in billiards a cue ball strikes a target ball there is no reason to expect any particular outcome; but this is to neglect the fact that neither is there any reason, in that case, to expect the cue to move the cue ball, or, another step back, to expect the player to be able to direct the cue, or ..., and so on. Hume could attempt to correct for this by saying that *even if* the cue ball strikes the target ball, no particular effect is to be expected, and this would at least draw attention to the scale of the threat that his arguments pose, not just with regard to what we know of the final effect in a causal sequence but to our knowledge of any part of it, and to the very idea that causal concepts have application. The point is that he does not do this, preferring instead to depict without question a man cueing a ball and straightening up to watch it collide with another, doubt intruding only with regard to what happens after impact. Hume thereby strengthens the assumption that the present reality of events now observed is outside the sceptic's field of fire.

I have argued against that assumption, but I shall now try to show, using our critique of Wittgenstein's *Investigations* as a guide, that the Humean sceptic cannot be consistent even on his own restricted terms, so that for the moment I need not appeal to his ultimately being committed to global scepticism. One way to show this is to return not to the snooker hall but to Hume's study, or that of a Humean sceptic. We may imagine that its occupant tries

very hard to conform his practice to his philosophical rejection of inductive inference. He succeeds, if only for a moment, in blanking out all knowledge of the world outside his study door; and then, epistemically cocooned, he contracts his purview even further when he closes the curtains on his window into the future; except, however, that he has to let in enough light to see the table and chairs in his study, which is to say to perceive them as presently existing physical objects.

But this, or so I claimed earlier, is to assume a cognizance of the physical world to which his inductive scepticism does not entitle him; and the reason is that his reference to physical objects or events goes beyond his immediate perceptual experience of them. He may now try to retrench at a more primitive level of perception, that at which sensory particulars are captured, according to Ayer, by intrinsic description. The objection here, however, is that intentionality is dispositional, a consequence of which is that even the most basic properties cannot be predicated of a subject, however attenuated one's reference to it, without implying knowledge of the future. Thus, if I say that this is red, succeeding in my reference only if I understand 'red', then it has to be that I am suitably disposed in my use of that word, which I imply would be in accordance with correct usage. But the Humean sceptic cannot imply this, for he claims to know nothing of the future, where this includes the conditional future, at least if previous arguments to that effect are not mistaken.

What this indicates, surely, is that there is something very wrong with a sceptical argument so pervasive in its ramifications that it prohibits any description, at any level, of the presently observed physical world. And the question it raises is that of whether Humean scepticism is self-refuting, given that *all* declarative use of language, as well as being occurrent, is also dispositional. In answering this question, let us begin by summing up the sceptical position in a single sentence: 'There is no reason to accept any empirical proposition about future events or unobserved present events'. This sentence is either understood or it is not. If not, then there is no sceptical position and nothing to discuss. If it is, then the sceptic implies that, other things being equal, he is able to use the constituent words in a variety of contexts or different words in a similar context; that, for instance, he knows the difference in meaning between 'there is no reason…' and 'there is always a reason…'. Also, he implies that he is able to recognise different instances of the same words, and in general that he is suitably disposed in a variety of ways in which he implies knowledge of what his responses would be. If our previous arguments were correct, the sceptic has no choice but to imply these abilities; yet on his own showing, taking his thesis to its logical extreme, he has no

reason to believe that he is able to do anything at all, let alone exhibit linguistic skills, including the recognition of linguistic signs.

This, it seems to me, poses a serious difficulty for the Humean sceptic, and my present aim is to multiply such difficulties in order to force him to admit defeat, this to be accomplished by our now capitalising on the fact that his arguments commit him to global scepticism. The first point is that as a sceptic about all empirical knowledge, he cannot even begin to be consistent, not even by closing the curtains, literally this time, on any windows in his study, or by immuring himself within. The reason is that belief about the present contents of a study, or of this room in which I write, is just as vulnerable to the global sceptic's arguments as any statement about objects that are spatio-temporally more distant. If this is difficult to accept it is perhaps partly because these walls enclose and indeed exhibit a great deal of visual stability, which lends structure to the more changeable aspects of what I see. This is, if anything, a psychological observation; it presupposes that the walls and furniture continue to look the same, my awareness of which depends on memory, recognition and expectation; and it may help to explain any resistance to the idea that the events of a moment ago, even if nothing much happened, are just as out of reach for the sceptic as yesterday's breakfast, the same being true of the solidity of this table and chair and these walls. This is hard to accept, and we may look to psychology; but it is *also* true, or it is an underlying truth, that one cannot doubt this present reality, for to make the attempt is to begin to lose one's grip on meaning and understanding before sanity is restored. This should bring the sceptic to a halt, but if he continues to nip at our heels then it just is a fact that there is no exclusion zone for scepticism, and certainly not in terms of spatial and temporal proximity to the here and now. To insist on this point is to require of the sceptic that he be consistent and accept his own logic, the consequence of which, or so I shall now maintain, is that his arguments refute themselves.

Let us begin by asking the global sceptic to modify the slogan already used, so that it now reads, 'there is no reason to believe any empirical proposition', which we may refer to as statement B. Then the difficulty faced by the Humean sceptic with regard to implying knowledge of the future is also confronted by the global sceptic, as I shall now show by elaborating on the previous arguments against the Humean sceptic. In asserting statement B, the global sceptic implies cognitive skills, an example of which would be his ability to understand the word 'believe'. This is shown by the fact that he would be inconsistent if he said, 'I have no reason to believe that I know what it is to believe.' In agreeing that he does understand and recognise that word, he implies restrictions on a possible future state of

affairs in which he again encounters or uses it, the restrictions arising from his conditionally expecting, as it were, to behave in ways appropriate to his present understanding, which is partly dispositional.

To go a little deeper, let us now focus on recognition, given that it is a necessary condition of understanding the word 'believe' that one should be able to recognise it. Then it is true that in agreeing that if the word 'believe' re-appears he will be able to recognise it, the sceptic means that he will be able to recognise it, other things being equal; and these include his not having succumbed to dementia in the meantime, this being just one out of a whole range of exclusion clauses. But there has to be a limit to what is excluded, otherwise to say that I can recognise a word again, other things being equal, is just to say that I can recognise it in all cases except those in which I can't.

There is also another argument which may be deployed against the sceptic, this time in terms of the anticipatory nature of declarative utterances, including his own sentence to the effect that there is no reason to believe any empirical statement. In uttering the first part of the sentence, he anticipates his next words, so that he knows what he is going to say, at least when this is opposed to his having no idea, for then it would not matter what the first part of the sentence was, or whether he remembered it; but it does matter, as does the way in which the sentence continues, and the difference it makes is both semantic and syntactical.

Let us now turn to the sceptic challenging statements about past events as part of his global scepticism. Clearly, he must imply that he is able to remember the first part of statement B; otherwise, he admits that he does not know whether he said, 'there is no reason…' or, 'there is always a reason…' Yet he is committed to the claim that he has no reason to believe anything about the past, thereby ruling out any knowledge of his own words, whether written, spoken or thought.

Objections

That, in outline, is the case against the sceptic before deeper issues are considered and more emphasis placed on the notion of a system, as it will be when we now examine objections to the arguments just advanced. Suppose, to begin with, that it is conceded that understanding is dispositional as well as occurrent, but with the caveat that one need not be *aware* of one's dispositions, any more than is the case with non-human creatures that exhibit a form of understanding. Then the reply should be that this is to reduce human understanding to the primitive non-verbal level at which carpet mites modify their conditioned responses when kept as pets. But philosophy is language-based and uniquely human, a truism that not even a

philosophical sceptic could dispute. If I say that I believe that I can count to 100, and if I am asked whether I would be confident in reciting the numbers, to which I reply that I have no idea, then this indicates that I do not understand what I have said. Similarly, if I say that there is no reason to believe any empirical proposition, and if I am asked whether I am suitably disposed with regard to my use of 'believe', to which again I reply that I have no idea, then again the suspicion will arise that I do not understand what I said. Carpet mites do not claim to understand anything, but people do, and on that basis if I claim to understand a word, where such a claim is implicit in my declarative use of language, my having particular dispositions being a truth-condition of understanding, then I imply that I am suitably disposed. This, however, is just what I cannot coherently do if I am a global sceptic, and also, as we have seen, if I am a Humean sceptic, for I should have to give an inductive answer if asked how I know that I have the requisite dispositions. I could, of course, profess ignorance as to how I know, but this is to admit that I have no evidence for my belief, and it can hardly be claimed on that account that my belief is impervious to the arguments of the sceptic; on the contrary, it is, if anything, even more vulnerable to attack.

Now suppose that in attempting to be consistent, the sceptic claims that although he understands proposition B, and believes it to be true, so that it is a statement he asserts, he has no reason, given its content, to believe that he will continue to understand it from one moment to the next. Then the rejoinder this time is that he misconstrues the role of dispositions. As a global sceptic, or just a sceptic about induction, he has no reason to believe that he will understand *anything* a moment from now, for he cannot expect to continue to exist or to keep his wits about him; but if his point can be made at all, then he has to mean that there is no reason to believe, with regard to any particular disposition of his, that the counterfactual conditional by which it is expressed is true, as opposed to the truth of the antecedent not being coupled with that of the consequent, so that the conditional is false. This, however, conflicts with his claim to understand statement B. Given, for instance, that it is a condition of his grasp of the word 'believe' in that statement that he would recognise it again, other things being equal, he cannot deny that he would recognise it; even with the ceteris paribus clause, indeed as following on from it, he must agree that there are circumstances in which he would recognise the word 'believe', otherwise the conditional fails to have meaning.

In barricading himself against attack, the sceptic seeks to exploit a difficulty of analysis. I grasp the sense of the word 'believe' if, as a necessary condition, I would recognise it again, other things being equal, which is to say in normal circumstances. Set

against this, or so it seems, is the possibility that a moment from now I cease to understand that word, no longer being able to recognise it. But this, perhaps, is a cue for the notion of a system to make its entrance again, so that the role of the ceteris paribus clause may be spotlighted. Suppose, then, that I understand the word 'believe', so that I would recognise it again, other things being equal, and that a moment from now I will not understand it, and not be able to recognise it. Then the circumstances must be such as to leave undisturbed, retrospectively, my present understanding. In the context of a system I must be able to distinguish between my not understanding and my ceasing to understand. It is hard to imagine normal circumstances in which I judge that i have lost my understanding of a moment ago, as opposed, for instance, to the special circumstance of my suffering a sudden stroke – which the ceteris paribus clause would accommodate. What this shows is that the sceptic's defence of his position does not hold up, for it really is the case that constraining future possibilities by way of conditional clauses is intrinsic to being suitably disposed.

Perhaps we should now defend and expand the earlier argument that knowledge of the past is a necessary condition of any declarative use of language, so that scepticism about such knowledge is self-refuting. Is it really true that I cannot state anything to be the case without remembering it? Clearly not, unless it is qualified in some way, for I am not required to remember what I said yesterday in order to have understood it. I shall now argue, however, by focussing on discourse rather than isolated utterance, that memory must come into play for a minimal period, albeit unspecifiable, if what we connectedly say is to make sense, so that we may understand it. Suppose I say that it is good to eat porridge because it enters the bloodstream and removes cholesterol from arteries by sticking to it and dislodging it, whereupon a congratulatory message is sent to the brain which induces euphoria, itself associated with the addictive properties of porridge and withdrawal pangs when it is not available. Now imagine trying to make sense of this final clause when you have forgotten the previous part of the sentence or attempt to disregard it in the interests of scepticism about the past. Not only does it not make sense out of sentential context but the sceptic could not in any case claim to remember it, as would also be true of his utterance of a complete sentence, however short, such as 'Porridge is delicious.' If, despite what he seems to remember saying, he has no reason to believe that the first word was 'porridge', so that he has to disregard his apparent memory, then there can be no question of his understanding what he said. Perhaps he said 'roadkill is delicious' or 'porridge is disgusting', except that he cannot entertain these possibilities without the aid of memory, which is to say that he necessarily refutes himself.

Further development of the solution

I think that this is enough, overall, to constitute a refutation of the sceptic, but to vanquish him completely we need to show, or show more clearly, that the beliefs he is obliged to reject are woven into the fabric of human discourse, including philosophical debate, so that he is not coherent in denying their legitimacy. To that end, consider my claim at time *t* to understand the even number sequence, and suppose that I substantiate it, or so it would seem, by reciting consecutively the first one hundred terms, at which point I meet not with applause but with its being denied that I have made good my claim, and for the reason that it was subsequent to *t* that I manifested my understanding, which it is therefore possible that I acquired after time *t*. If that is what a sceptic would say, then his thesis is that criteria of understanding have application only to what it is that one now understands, at the time that evidence in the form of cognitive behaviour is forthcoming. The fact is, of course, that it is the thesis itself which cannot be applied, and for the simple reason that cognitive behaviour has duration, so that nothing could count, if the sceptic is correct, as one's present grasp of, for instance, the even number sequence. If I start to recite the initial terms at a particular moment, then immediately I am swept beyond it to the next, so that I call out the second term, strange to say, after I have called out the first. What this shows, the sceptic notwithstanding, is that intention concepts are *transtemporal*, which in the present case is to say that if at time *t* I claimed to understand the sequence, the initial terms of which I 'now' call out, then I am judged to have understood it at time *t*, other things being equal, and to *continue* to understand.

This apparent underdetermination of understanding by evidence is not a contingent fact but a fundamental condition of discourse about belief, understanding, and so on, one that is so familiar to us in the way it plays out, being intrinsic to belief and reasoning, that we need to focus on it in order to bring it clearly into view. What is then revealed is not just that my cognitive behaviour inductively indicates to myself and others that I am suitably disposed, but also that factual inference, which includes induction but goes wider and deeper, necessarily enters into what it is to mean and to understand. At this deeper level memory permeates induction and in particular cases may be justified in terms of it; but also, in terms of its place within a system in which it cannot be doubted, either in particular cases or in general, for any reason that also threatens the expression of that doubt. The appeal to self-reference in the context of a system, then, takes us deeper than induction does into the ways in which language and reasoning are connected by way of criteria of meaning and understanding. Much the same may be said, with suitable adjustments, about belief avowals

and claims to be suitably disposed, a topic which we have already examined in considerable detail. It has been argued, too, and in relation to what it is to understand, that the sceptic about understanding has to avoid self-refutation, which is to say that it is only in particular cases that doubt about one's own understanding may coherently be expressed. There is also the fact, which I could have made more of, that all these structural components enable the operation not only of factual belief, including induction, but also of logical and mathematical reasoning, which provides a touchstone by which factual inference is deemed by the sceptic to be deficient – and this is a fact the anti-sceptical significance of which I shall not now investigate.

Interdependence of past, present and future

There is, however, a related investigation I would like to pursue, given that I have just compressed into one or two paragraphs the main conclusions of the transcendental arguments advanced in these chapters against both the empiricist and the sceptic. The thesis for which I have argued is that factual discourse has a deep structure, in some ways hidden and in all ways impervious to the sceptic, such that reasoning about the world operates within a system, one that is characterised by interconnectedness. Needing to be examined, then, is the nature of the connections, and what I shall try to show, already implicit in many of the lines of argument already followed, is that the various parts of the structure are interdependent. As a step in that direction, one may ask whether memory and expectation, since they are intrinsic to every declarative use of language, are necessarily bound together. One way to do this is to allow the sceptic to try to target them separately, expectation or prediction being considered first, so that we may ask whether it is conceivable that one should be sceptical about the future but not the past, this time from the point of view of a sceptic who simply chooses, and without making a case, to disregard all reference to future events. Perhaps this could be taken both at face value and as a device by which to spotlight the mutual dependence of the concepts of memory and expectation.

Suppose, then, that the sceptic wakes in the night, opens his eyes, sees a yellow disc shining in the dark and, understandably, closes them again. He says, 'I saw a yellow disc.', but as a sceptic about the future he also says, 'I have no reason to believe that I shall see or remember it again.' Now he opens his eyes, finds that it is still there, and says, 'I see a yellow disc, the same as before.' Then the first point is that he has no reason to believe that he remembers anything. In the normal world I see a disc, close my eyes, expect to see a disc, open my eyes and see one. This confirms my memory, and in the strongest possible terms, for

how else is it to be explained, courtesy of the no-miracles argument, that I correctly expected to see a disc?

In the world of the selective sceptic, however, nothing could count as confirming his memory or, indeed, any of his beliefs. Well, could it not be argued that even if he did not expect to see a yellow disc, or anything else, but did see one, then this confirms his memory of the previous observation? If we take the disc to be a solid object, after all, then such objects have identity through time, other things being equal. But now, the sceptic cannot avail himself of such facts, for he has no conception of a physical object, since this would involve his expecting its appearance to remain the same or to change, depending on the circumstances; and, therefore, he cannot conceive of a physical world which exhibits any form of continuity. It is arguable, even, that he is not able to identify perceptual objects, for instance what he now sees glowing in the dark, as coming under a particular description, even in terms of sense-qualia, or as being similar to a previous object, in this case one that he has just seen; and the reason, familiar to us from an earlier discussion, is that to identify or re-identify particular items is to imply constraints on future possibilities. If this is correct, then in what sense, given these epistemic and semantic limitations, his words not having their conventional meaning, does he remember having seen a similar object to this one? With regard to confirming his memory of the disc, he might as well, surely, have opened his eyes and gazed upon a luminous yellow crescent – as we, but not he, would be able to call it, just as we are able to recognise it as a crescent moon.

We may argue in the same vein if we now shift from scepticism about the future to that about the past. Suppose that once again the sceptic, this time a sceptic about memory, sees a yellow disc, at which point he closes his eyes and predicts that he will see it when he opens them. Then it is immediately apparent that the situation we are trying to imagine lacks all intelligibility. If the sceptic opens his eyes and sees a disc, then nothing is confirmed, for he cannot confirm a prediction that, for the purpose of this discussion, he does not allow himself to remember making, just as he does not remember the beginning of what he says by the time he gets to the end. Thus it is that the concepts of memory and expectation are intelligible only when linked together, a fact which indicates that a relation of mutual dependence may hold between intention concepts in general, a consequence of their interconnectedness.

Resolving the sceptical hypothesis problem

One could go on, and no doubt there is more work to be done with regard to consolidating this thesis, but what I now propose is to place it within the contemporary debate on epistemological scepticism, hoping thereby both to test it against sceptical arguments not yet considered and, conversely, to confront those arguments with the insights gained from it. There is, for instance, what has become known as the brain-in-a-vat problem, and it starts from the sceptical hypothesis that we are disembodied brains programmed with those experiences on which we base our apparent knowledge of the external world. It would seem, on the face of it, that physical theory and the sceptical hypothesis are both compatible with the experiential evidence, a question then arising as to how we can claim to know that the world exists, and in particular that our brains are housed in bodies, not vats, so that, for instance, we have hands. This is a variant of Cartesian scepticism, and schematically the argument may be set out as follows:

1. I do not know that not-H.

2. If I do not know that not-H, then I do not know that O.

So,

3. I do not know that O.

Here the sceptical hypothesis, H, might be that I am a brain in a vat; the everyday fact, O, that I have hands, which is incompatible with H. This argument seems to depend on the principle of epistemic closure, such that if I know that O, and know that O entails not-H, then I know that not-H. By contraposition, if I do not know that not-H, then I do not know that O. And the reason I do not know that not-H, where H is the brain-in-a-vat hypothesis, is that this latter is consistent with all my external world evidence, for if H is correct then the course of my experience will be just as it now is.

If the point of the argument is merely to demonstrate that I cannot infallibly know what I claim to know, for instance that I have hands, given that it is logically possible that I do not, then it may occur to us to concede the point without feeling that we are sacrificing our everyday certainties, which would still be justified, at least for all that has been shown to the contrary. Michael Williams (1996, ch.2) agrees, and he suggests, indeed, that if the hypotheses in question are taken to support radical scepticism, then they can offer that support only if foundationalism is presupposed in the sceptical argument. Be that as it may, the assumption seems to be that the argument threatens our everyday knowledge, and then a question arises as to whether there is a general principle or counter-argument by which our practice of dismissing such deviant hypotheses may be shown to be vindicated. In search of

an answer, I propose to consider a particular form of contextualist rejoinder to the sceptic, which proceeds from the view that we judge our beliefs against a background of relevant alternatives, using criteria of relevance by which alternatives that are no more than logically possible are normally, but not always, excluded. A familiar example, first given by Dretske (1999), is that of claiming to know that an animal in a cage is a zebra, in which case the sceptical hypothesis might be that the animal is a mule painted to look like one, the argument running as follows: My evidence that this is a zebra would be just the same if it were a mule; therefore, I do not know that it is not a mule; therefore, I do not know that it is a zebra. But there seems to be a paradox here, for I *do* know that it is a zebra.

Let us now see if the relevant alternatives view offers a resolution of the paradox. On that view, according to DeRose (1999), the painted mule hypothesis becomes relevant not only in easily envisaged special circumstances such as an outbreak of mule painting being reported, but also in a philosophical context in which a sceptic challenges our claims to knowledge by invoking sceptical hypotheses consistent with the evidence on which we base our everyday beliefs. In that context, it will be said, the sceptic is correct in claiming that we do not know that we have hands, or that there are zebras, and our having to agree with him may seem to us to conflict with the fact that in ordinary circumstances we *do* have this knowledge. But that, according to the contextualist, is where we go wrong, for if knowledge is context dependent then there is no conflict. Ordinarily we know that we have hands; it is just that if a sceptic raises the epistemic standards by positing that we are brains in a vat, then in that context we do not know that we have hands, or not unless we can rule out the sceptical hypothesis.

But now, this contextualist thesis seems counter-intuitive, for we are certain that we are not brains in a vat, and that conviction seems to carry over from everyday situations in which the hypothesis merits no consideration, and for the reason that there is not the slightest evidence in its favour, never mind the very strong evidence that taking it seriously would demand. Even with regard to identifying zebras, it surely is not the case that mentioning the painted mule hypothesis would be enough in itself to make it relevant, as opposed to adducing evidence that elevates it from the level of logical possibility to that of serious alternative. Besides, the proposed solution, like Kripke's to a different problem, would seem to be tantamount to the scepticism it is meant to fend off.

It is worth asking whether an enquiry into the nature of reasoning might yield principles by which we take a given body of evidence to point to a particular conclusion, which it underdetermines, rather than to other conclusions to which it seems to be equally

disposed. And this, as I shall now try to show, is where the notion of a system comes into its own. If I visit a zoo and identify an animal as a zebra on the basis of observing it, then this involves perceptual and cognitive processes which operate within a system, and it is one about which, in a quite peculiar sense, I know nothing. I can say that my being able to identify zoo animals depends on previous learning, details of which I may be able to provide; and I can envisage epistemic circumstances in which the painted mule hypothesis should be taken seriously – but all this is itself a matter of perceptual and cognitive processes, including those of memory, so that it occurs within a system. I know nothing of that system and have no control over it, for I feel myself swept from one belief to another by a wholly mysterious movement of forces, these being manifested only in experience itself, so that nothing counts as my peering down through the water to the source of it all. In a particular case one can say particular things, or one can generalise to some extent, as when I say that at a zoo there is never any reason to doubt one's identification of zebras. But this does not yield any general principles by which it is deemed that some alternatives are relevant and others not, or none that belong to philosophy rather than the natural sciences or the psychology of reasoning.

That said, the good news is that we can make effective use of the notion of a system in the same way as in other chapters, which is to say by enquiring into the necessary conditions of discourse and reasoning and by drawing conclusions about what it is possible to coherently assert. Suppose that I identify a zebra by its appearance and that a doubt intrudes when I take seriously a sceptic who points out that if the animal was a painted mule its appearance would be the same. In that case I am foolish on one count and ultimately incoherent on another. To take the hypothesis seriously is to forget that my observations are made against a background of knowing that animals in a zebra cage are almost always either zebras or keepers, easily told apart, and that it is unheard of for one species of animal in a zoo to be disguised as another in order to deceive the public. By dint of *probability*, then, there is every reason to think that this is a zebra and none to think that it is not. That is one answer, and another concerns the hidden incoherence of giving epistemic weight to negligible possibilities. If I allow the possibility of the animal being a painted mule to cast doubt on its being a zebra, then I commit myself to *generalising* this epistemic approach. But now, if I start from the belief that this is a zebra and then give epistemic weight to its possibly being a mule, then I am committed to weighing in the balance the contraries of all the statements to which I give assent, for all beliefs are underdetermined by the evidence, itself underdetermined, or they involve understanding, which goes beyond what is given, since nothing is just given, or they are taken, in a sense, to stand on their own. It makes no

difference which it is, for the result in any case would be incoherence and cognitive paralysis. If my belief that this is a zebra is infected with doubt that feeds on logical possibility, then the virus will attack the beliefs that enter into that belief or into the possibility that the animal is a painted mule, or that it is a hallucination, my cup of tea in the zoo cafeteria having been doped, or that it is a zoo keeper whom I see as a zebra, the real zebra appearing as a zoo keeper, the one brushing the other's tail. And all these possibilities in their turn would be no sooner entertained than doubted and set against their contraries, and so on...Nothing counts as a viable proliferation of doubt if logical possibility is the vector, for the body of belief and evidence would be lifeless in a former sea of discourse from which all meaning had been eliminated. Clearly, the underdetermination of belief by evidence is *fundamental* to one's cognitive and perceptual experience of the world.

How, then, is the original problem to be solved? The sceptical argument given earlier in schematic form starts from the premiss that I do not know that I am not a brain in a vat, for if I were then my experiences would be the same. If the discussion thus far is along the right lines, then this first premiss is mistaken in what it asserts, so that the problem is thereby resolved. From the fact that any particular knowledge claim is always based on evidence with which at least one sceptical hypothesis is consistent, it follows not at all that the claim should be revised, for logical possibility cannot be a basis for legitimate doubt, and it is a fundamental condition of discourse that all empirical beliefs should have logically possible contraries. If it is now asked whether infallibility is intrinsic to the concept of knowledge, the question is, I am sure, not without interest, if only for the reason that to answer in the affirmative is to imply, rather contentiously, that the verb 'to know' is invariably misused when positively applied, for it is always possible that we are wrong when we think that we are right. In the present context, however, that question may be bypassed, for it is open to us, inconvenience apart, to speak not of knowledge but of justified belief and entitlement to certainty – or virtual certainty if one is going to be pedantic, or perhaps overly pedantic, it being arguable, after all, that logical possibility in itself has *no* epistemic weight. If that is the position taken, then it has to be squared with the fact that some of the various possibilities we have considered feel epistemically heavier than others That we are brains in a vat is a mere logical possibility and lies at one extreme on the epistemic scale, whereas a supposed zebra being in a particular case a cleverly disguised mule is more easily accommodated in one's thoughts, for the possibility, if realised, would do far less violence to one's preconceptions. Perhaps better, then, to place logical and negligible possibility in different categories, the second included in the first; but for present purposes it does not matter, for the brain in a vat

hypothesis is only a logical possibility, and the painted mule hypothesis has negligible weight and lands like a feather, as it were, on one's identifying of zebras.

Wittgenstein's *On Certainty*

Now that we have resolved the brain-in-a-vat problem, let us compare our use of the notion of a system with that of Wittgenstein in *On Certainty*. According to Grayling (2008), there are two conflicting themes in the book, one being a form of relativism and the other a response to scepticism which combines foundationalism with appeal to the notion of a system or framework. If the themes conflict, says Grayling, that is because Wittgenstein's relativism weakens the anti-sceptical thrust of his development of the notion of a system, and of a system having foundations. He speaks, for instance, of knowing that the earth is very old and that no-one, at that time, had travelled far away from it, these being certainties that he shares with Moore and that he regarded as part of one's picture of the world and as having foundational status. (411) But then he speaks of this picture as being mythological (95), and he goes on, 'The mythology may change back into a state of flux, the river-bed of thoughts may shift.'(97) Grayling gives these and other examples from *On Certainty*; and the reference to a state of flux does seem to detract from any anti-sceptical potential of drawing attention to a solid core of established beliefs. But also, even on its own relativistic terms it is problematic, for if the age of the earth had to be radically revised then this would count not as a change of framework or conceptual shift but as a massive intellectual earthquake shaking to its foundations the whole of scientific theory and methodology, one that could be brought about only by the most powerful evidential forces that – but then at this point one's imagination fails, for what could possibly make it reasonable to believe that the earth is only a few thousand years old?

Since I agree with Grayling about the conflict in *On Certainty*, and since its relativistic theme resonates not at all with the views developed in this thesis, what I propose is to focus on those aspects of Wittgenstein's thinking that may have affinities with my own approach. Immediately, however, we encounter a problem, for when Wittgenstein makes play with the idea of foundations underpinning the system, he conflates contingent belief with what he refers to as facts of grammar, a tendency which Grayling associates with his relativism. When he says, as he does, that we know that there are physical objects, the way in which this is a grammatical fact is clear enough, for any attempt to *argue* that there are physical objects would have to presuppose them, the existence of the physical world being the framework in which we acquire particular empirical beliefs. But in the same breath he

speaks of the age of the earth, and to believe that it has existed for aeons is to partake of scientific wealth amassed over the last few hundred years, the age of the earth being contingent in a way that the existence of the physical world is not. The difference, of course, is that I spend every moment of my waking life within a cognitive and perceptual environment in which the physical world is always insistently *there*, whereas the earth's antiquity is just another incontrovertible fact of science.

The key question, anyway, is that of whether Wittgenstein's response to the sceptic is effective in refuting him. The radical narrow sceptic about induction would agree with him, and with Moore, that the earth is old and that this is a hand, for he is unaware of being committed to global scepticism; but he would challenge them to show that their hands will continue to exist, or that probably they will, at least from one day to the next, which no doubt they would also claim to know; and I have argued, indeed, that present-tense observation reports imply constraints on the future. But the fact is that the appeal to the notion of a system in which the kind of prediction we take for granted stands fast, or is foundational, or in which prediction in general is part of the framework, as Wittgenstein would have to insist that it is, would fail to impress the inductive sceptic, who would point out that Wittgenstein is begging the question; and this is a charge which the global sceptic, or the fully committed inductive sceptic, would level at him with respect, also, to his claim to have knowledge of the past.

Wittgenstein does not, in my view, develop the notion of a system in such a way as to pose a threat to the sceptic, and I do not see how this can be done except by showing that inductive scepticism is self-refuting, as also the global scepticism which it entails. To point out, as he has done, that our cognitive and perceptual interactions with the physical world form a system is to take only the first step, except that even as a first step it seems to be in the wrong direction, for he speaks not only of a system but also of foundations, as with our knowledge of the earth's antiquity; and this is to appeal to a core belief embedded in a rock of established theory which itself is deeply entrenched in layers of discovery and experiment accumulated over time. But this appeal to what seems to be so solid is, as I have said, ineffective against epistemological scepticism, for what is needed is not that we cling to the rock but that we combat the sceptic, Hume in particular, by refuting the arguments by which he contends that the rock is hollow.

Inductive inference reconsidered

This is a project which, as I have shown, can be carried through only by penetrating deep into the nature of belief and understanding and extracting the fundamental conditions of

discourse. What is fundamental is not that the earth is old or that it rained yesterday or that a moment ago I held up my hand, but that reliance on memory is an *absolute* condition of knowing, believing or doubting anything at all, as is reliance on expectation and prediction, together with dispositions being sacrosanct, beliefs being underdetermined, and so on. Once this has been shown and the sceptic's vulnerability exposed, so that he cannot help but refute himself, we are then free to develop further the notion of a system and to apply it to other areas of interest, as with the problem of other minds, to be addressed in the next chapter. In this chapter, as part of that development, I have attempted to give sense to its being said that we just *have* beliefs and that we are carried from one inferential step to another. Clearly, there are insights to be gained from ploughing this furrow deeper, now that the sceptic lies dead in the next field, so that there is no longer a risk of being charged with psychologism or being taken to dissociate evidence from belief and to imply that there are no grounds for what we believe.

It remains the case, then, that I expect the cows to come home because they always have, a question arising only about the place of regularity within an explanatory system. We return, then, to the question we left with Okasha and Lange, the impetus for which comes partly from our undermining of the distinction between inductive and global scepticism, suggestive as it is of induction not being fundamental in the scheme of things, that status being awarded to necessary conditions of discourse. Accordingly, what I now propose to do in the remainder of this chapter, apart from preparing for the final one, on other minds, is to bring our campaign against the sceptic to a close by re-assessing the role of induction in factual reasoning. This will involve reaping what was sown in our treatment of intentionality in previous chapters, and perhaps being able to harvest a contribution to confirmation theory. To that end, we need to focus more on the nature of induction than on justifying it, and in particular on what it is for a predicate to be projectible, given that past instances of some properties or events, but not others, are susceptible of extrapolation. There is, too, a related question as to the principles by which instances are identified and labelled, and as to the link between language and the locating of similarity and difference in the world. This is all very abstract, so perhaps we may give substance to it, as always, by fleshing it out in the form of an example, in this case one that I hope will dispel darkness in favour of light.

Let it be supposed, then, that upon entering a room I flick a wall switch and expect a light to come on. For my expectation to be met, it will have to be that the electricity is on, that the switch and bulb are connected, that the bulb is working, that switching it on will not break it, ..., and so on. These are empirical conditions for the correctness of what I expect, but

I need not be aware of them, or of all of them, in order for my expectation, in some form, simply to obtain. If I *am* aware, for instance that an electric current needs to flow through the filament in the bulb, then my expectation has to epistemically accommodate that awareness, for it would make no sense for me to say that I expect the bulb to light up, knowing that this requires electricity, but that I have no idea whether it is electrically connected. If I am totally unfamiliar with the electrical world, 'electricity' and its cognates being absent from my lexicon, then it must be that by 'light switch' and 'light bulb' I mean less than other people do. The point is just that however circumscribed my knowledge, provided I am a cut above carpet mites, it should always be possible for me to establish a causal connection based on past instances of pressing a switch and seeing a room light up. At the level of philosophical discourse this lends itself to the view, which I hope to be able to challenge, that simple cases of inductive belief, as with the present example, exhibit universal features of induction which characterise even the more complex cases. What they all have in common, it will be said, is that the repeated instantiation of a projectible property or event generates and justifies the expectation of further instances. They are instances of narrow inductive inference, defined earlier as being linear, epistemically separate from any background information, and based solely on repeated instances which determine the nature and strength of the resultant belief. On this view, then, there is no reason why we should not focus on simple cases, as with pressing a switch in order to turn a light on or off, and we may also make use of the fact that basic objects of perception, such as colour and shape, are projectible when predicated of stable physical objects. It is possible, for instance, to generalise from a sample of emeralds, each of them observed to be green, to the conclusion that greenness is characteristic of all emeralds. What this may imply, in the context of confirmation theory, is that the greenness of observed emeralds has intrinsic inductive value, given that Hume's problem may be ignored, perhaps because it is taken to have been resolved. That inductive evidence is in some sense intrinsic is a proposition with which we are already familiar from the chapters on Fisher and Keynes, this latter arguing that probability is relative to evidence and that it is a logical relation between evidence and conclusion. Here, too, it is implied that probability evidence is self-contained and has its own epistemic value, even if this may change when further evidence is considered. Set against this is the notion of a system, which I developed as a rebuttal to the attempt by Keynes and others to bridge the gap between induction and logic. Already in the armoury, then, are arguments which run counter to the view that intrinsic evidential relations obtain between the instances which enter into the premisses of an inductive inference.

Goodman's paradox

What I now suggest, before deploying those arguments once again, is that we consider a very different kind of challenge to the view in question, one familiar to gemmologists in the form of Goodman's paradox or the new riddle of induction. If we enquire into the distinctive features of emeralds, taking greenness to be a contingent property, not part of the definition, then we reason inductively when we say that all emeralds are green. Goodman does not dispute the rationality of this, but he poses a question as to what it is about properties such as greenness that makes them projectible, an answer in terms of natural kinds not being found adequate. To expose the difficulty, using as an example the hypothesis that all emeralds are green, he constructs the rival hypothesis that they are grue, an object being grue if it is first observed before time *t* and is green, or after *t* and is blue. They compete, he says, because emeralds observed before *t* to be green are also grue, so that such instances equally confirm the conflicting hypotheses that they will be green after *t*, and that they will be blue. None of this, as it stands, is of striking interest, for it seems clear that 'grue' is parasitic upon 'green' and 'blue' and that it is unprojectible, being temporally positional in its disjoining of the two colours. What Goodman now does, however, is to introduce the term 'bleen', such that bleenness is a property of emeralds observed before *t* and found to be blue, or after *t* and found to be green. He is now able to take these new predicates as primitive, so that 'green' and 'blue' are defined in terms of them and thereby become positional, an object being green if it is found before *t* to be grue, or after *t* to be bleen, and the other way around for an object being blue. Goodman now asks why we should take green and blue to be primitive, as opposed to grue and bleen, the point being that no answer is readily available, where this is taken to illustrate the general difficulty of accounting for the way in which we pick out some predicates but not others as projectible. The answer that he eventually settles on is that the predicates which we find it natural to project are those that are entrenched in the language. This attempted resolution of what is claimed to be a new problem of induction need not detain us if it can be shown, as I think it can, that the seemingly paradoxical nature of Goodman's invented predicates dissolves under scrutiny.

To begin with, a paradox arises only if 'grue' and 'bleen' are read as colour predicates, so that they refer to qualitative observational properties of objects; otherwise, for instance if they referred to shape, they could not clash with 'green' and 'blue' by way of conflicting predictions of the colour of emeralds after time *t*. What this conflict also requires, if our usual understanding of induction is to be put out of joint, is that the rival predictions should each follow inductively from the evidence, which is to say from the fact of emeralds

examined before *t* being green, in other words grue. This is possible, according to Goodman, if 'grue' and 'bleen' are treated as primitive; and it may then be said that emeralds unearthed after *t* are inductively predicted, before *t*, to be grue, the same as their predecessors, so that there is no change of colour. What is paradoxical is that for us the emeralds would then be blue. But now, it just is a fact that these deviant predicates *are* positional, not primitive, even if they are interdefinable with 'green' and 'blue'. If we imagine that a child is taught their use, but perhaps not that of 'green' and 'blue', then he can distinguish between a grue and a bleen emerald only by knowing whether the time is before or after *t*, and whether, if after, the emerald was examined before *t*. The appeal to interdefinability should therefore not detract from one's natural response when first encountering Goodman's definitions: that 'grue' is a disjunctive colour predicate in virtue of which it is non-projectible, for instances of green emeralds before *t* can confirm only that all emeralds are green, not that those examined for the first time after *t* will be blue; and they can do this only if they are regarded as constituting a fair sample from which the colour of emeralds in general may be inferred.

This discussion of these deviant predicates has run along lines familiar in the literature[2], but what I now propose is to draw upon the work of previous chapters in taking a fresh look at what it is to reason inductively and at the role it plays in factual belief. As already noted, an assumption underlying Goodman's paradox is that induction depends on intrinsic evidential relations between similar instances of a suitable kind, the task being to ascertain the criteria by which some predicates, but not others, are deemed to be projectible. What I have in mind is not so much to undermine this assumption as to place it in context, though it may turn out to be much the same thing. If we return to ornithology and begin not with ravens, which are too often disturbed in their nests by philosophers, but with swans, then it may be supposed that on the basis of my observations I expect all swans to be white, given that I myself am not Australian. But now, I can expect this of them only if I am able to identify them as swans, even without their being white, and this entails knowing what they look like in a variety of conditions and from varying perspectives. If I am asked to justify such knowledge, then I can appeal, more or less vaguely, to my past experience of swans. But it is quite likely that I cannot cite any particular past observations, and in any case it is difficult to distinguish between the source and the justification of inductive belief, not to mention the fact that if I happen to be in the vicinity of a swan, with a sceptic in tow, then all I have to do is to describe what the swan will look like from different angles and then walk or

---

[2] It is much indebted to, for instance, Simon Blackburn's treatment of the problem in *Reason and Prediction*. See bibliography.

swim around it in order to show that my description is correct, which I already believe it to be. This is enough to suggest that there may be a sense in which I just *have* the beliefs about the swan, and that in suitable circumstances my concern is only to exhibit their correctness. That beliefs are in a sense just there, that they come and go, as it were, so that they either obtain or they do not, has already been implied in the treatment in the previous chapter of belief avowals. If belief and expectation as occurrent and non-reductive entities are part of a cognitive and perceptual flow in which understanding is conveyed from one moment to the next, then this is to say that even when belief is inferential, clearly derived as a conclusion from given facts, the beliefs by which those facts are interpreted and given evidential force are themselves just part of what is given, so that they simply obtain, or they change under pressure from other beliefs which simply obtain; and this is to say that *all* beliefs, even those that form a reasoned conclusion, simply obtain, as does the reasoning itself.

What is also clear, even without the tracks laid down in previous chapters, is that justification depends on memory, in this case either of past observations or, more immediately, of one side of the swan from the other. I can support my claim to remember these things by appeal to my memory being reliable; but this itself depends on memory, which informs every conscious moment, so that it is inescapable. This gives the sense in which I just *have* memory-beliefs, which is not to say that no grounds can be found for them; on the contrary, they are self-intimating within a system in which, as we have seen, memory, expectation and confirmation are necessarily connected.

Continuing in this vein, if I see or feel a swan my registering of its particulars as a three-dimensional object goes beyond what is immediately given and is informed by my 'seeing' the swan as three-dimensional, an experience which I simply *have*. It, too, is self-intimating or intrinsically informative within a system, one in which the 'seeing' of the objects we see is essential to our locating them in physical space and to our visual knowledge of them. A similar point may be made about my self-ascription of beliefs about swans, where my claim to hold a particular belief is authoritative if I understand it and am not pretending, the focus thereby shifting to what it is to understand. We have seen that understanding is both occurrent and dispositional, and that it is only within a framework of understanding that we can think that we understand or that we do not. Our reasons for doubt cannot be such as to have universal application, but must apply only to particular cases. Understanding entails belief, without which nothing could be asserted or denied, and again there is a sense in which I just *have* a belief, in the context of the self-ascription of belief being inescapable.

All these essential components of reasoning and knowledge may be given an inductive gloss, as with memory being reliable, or with our mostly being correct in our claims to believe or to understand, or in our going beyond the phenomenology of particular physical objects when we identify them, and so on. Indeed, much was made in previous chapters of the all-pervasiveness of induction, which itself enters into the premisses of any inductive argument and into the descriptive use of language. But it is also true that those essential components, by which the necessary conditions of discourse are satisfied, themselves pervade induction. If this is accepted and its ramifications worked out, then it will seem quite natural, and will not cause any net curtains to twitch, that one's expectations about a particular object or event should be based on a single observation, or perhaps on none at all. If I open a box, glance into it and then close it, confident that I have glimpsed a swan inside, then this is to say that from a myriad of possibilities I have embraced a single one on the basis of a patch of white, a dash of black and orange and a beady eye. One could, as always, rationalise this process by appeal to induction, for it is arguable that observational beliefs, although underdetermined by the perceptually given, are usually correct. What counts, however, is the particular case, and the fact is that I looked into the box and just *had* the belief, as an integral part of my perceptual experience. It may not be without interest to note, too, that perhaps I have never before seen a swan in a box, or at least not in a telephone box on a Sunday.

But suppose, it might be said, that there are numerous boxes, not telephone boxes but proper ones, of which each of a random sample reveals a swan inside, prompting the expectation that at least one other will also contain a swan – is this not a paradigm case of inductive inference? Indeed it is, and we are rational if, other things being equal, we expect more swans; or if, other things not being equal, we do not, for we may already believe, perhaps from testimony, that the boxes contain a specific number of swans, these all being accounted for in those already opened. If that is the case, then we now have to weigh in the balance conflicting evidence, and of different kinds: testimony on the one hand and the no-miracles argument on the other. But the no-miracles argument stands on its own in the sense that we may employ it without realising that it is frequency-connected within a system. As for testimony, if I am told that there are *x* number of swans then I have to judge whether my informant is reliable, which quite possibly is a matter of the impression I have of her. Again, then, we are back with judgements which are far removed from paradigmatic cases of induction, and it is arguable that even in these cases we have an idealised view of the link between evidence and belief, taking it to be linear rather than branching out into a network. The metaphor of a network is more appropriate not only when evidence conflicts but in all

cases of factual inference, for it is always true that a variety of factors determine inferential connections. Thus, I see what I take to be the front of an old building; or, knowing differently, I take it to be the original facade of a building now replaced with a modern interior. Another example, one which returns us to Goodman, is of expecting emeralds to be green before a particular date, and blue if unearthed later; but this time one has to be imaginative to be able to conjure up circumstances in which such a belief could arise.[3]

In the context of this enquiry into the nature of induction within factual inference, we are now able to challenge the assumptions underlying Goodman's paradox, asking first of all whether there is such a thing as a projectible predicate or property. Well, if I glance inside a box and seem to see a swan, then something is being inferred on the basis of that glance, just as it would be, though perhaps less obviously, if I gave the occupant of the box a closer look. Still, the epistemic reality is that my identifying it as a swan is not in doubt – or, if it is, that at least I recognise it as a bird; or, at least as a living creature; or, if I have not watched any nature programmes at all, as a feathery physical object. The point is that I slot what I see into a pre-existing epistemic system or explanatory framework; and, that being the case, we may pose a question: where in all this are the predicates with regard to which there is a difficulty about ascertaining the principle upon which some are projectible and some not? But was it not said, albeit as a joke, that one may never before have seen a swan in a telephone box on a Sunday? Here, however, the question of projectibility yields to that of epistemic relevance, so that we may now ask why it is that location and day of the week are taken to be irrelevant to whether a particular object is a swan. But now, what could be *more* relevant than these two factors if, for instance, one is not initially sure whether the bird in the telephone box is a swan or a seagull, that uncertainty being dispelled when one then remembers that seagulls never make calls on a Sunday? In other words, the question of relevance is answered by facts about the world, and in particular cases, not by general facts about projectibility.

In conclusion and by way of summing up, if Goodman's paradox is directly addressed and found wanting, as happened earlier, then we are free to take a wider view and to show in a different way that the paradox is not genuine. If it seems to have substance, that is because it trades on our ordinary conception of physical objects, in this case emeralds, which it treats as retaining their specific and numerical identity over time. There is no new problem of induction, as our investigation of the old one has helped to show, and what it has revealed, compressed into a single proposition, is that nothing can be said that does not go beyond the

---

[3] Such circumstances are imagined in Blackburn's *Reason and Prediction*.

immediate moment, which therefore does not exist. This, however, is not a paradox but a mystery.

Finally, in shooting a quiverful of anti-sceptical arrows at Hume's problem, I have turned the inductive sceptic's arguments against himself, and in such a way as to combat scepticism on a broader front. Thus, it should now be safe to believe not only that the sun will rise tomorrow but that it has risen in the past and is presently aloft somewhere in the sky. This is good news for astronomers and careful sunbathers, but in philosophy a major department of our most cherished everyday beliefs still awaits reinstatement, having been suspended in response to a challenge by a particularly disturbing kind of scepticism: that which calls into question our knowledge of other minds. In reality, of course, our beliefs about one another continue in employment, and are put to work just as much by the philosopher as by the layman, as is obvious from the fact that philosophers would not be in relationships if they genuinely believed that their partners were automatons.

I shall argue in the next chapter that solipsism is a logically valid theory of one's place in the world; and it might even be said that its practitioners are scarce on the ground for evolutionary reasons: over the course of human history they have met with reproductive difficulties and not been able to propagate their point of view. That, at least, is a suggestion I once overheard in a public bar. At any rate, I shall try to show that solipsism is not self-refuting, which is to say that the foregoing anti-sceptical arguments have their limitations, this being one of the points I wish to make. By way of balance, however, I would add that there is much in the approach we have taken that is transferrable to this new problem, some of the arrows finding their new target, and that the solipsistic sceptic can expect strong opposition when he attempts, rather inconsistently, to convert us to his philosophy of mind.

Relaxing effect of problem resolution

That, as I said, is for the next chapter. To bring this one to a close, I would like to report that I now sleep easily in my bed, having resolved the induction problem at least to my own satisfaction. If I wake to a yellow globe glowing in the dark, I immediately recognise it as the moon, about which I am quite relaxed, so that I am soon asleep again. Sometimes, however, I reflect for a moment on the fact, which I am again able to take it to be, that the moon has shadowed the earth for aeons, as it will continue to do, rather annoyingly, after I am dead. Beginning to feel anxious, I now clutch at a special thought: that if the solipsist is correct, in which case he, she or it is mindless and a thing, since all consciousness in the universe must reside in myself, then I have no reason to believe that I am mortal, for I stand apart from all

other organisms. I now imagine myself as the ultimate Robinson Crusoe, a castaway on an island of the self surrounded by automatons, with Man Friday a comfort only in the way that a teddy bear would be. On this showing, then, one regards oneself either as living and dying among fellow human beings or as gazing forever across a sea of empty faces. To help us to decide, let us now turn to the final chapter of this thesis.

Chapter 7

The Problem of Other Minds

The sceptical argument

Since my aim in this final chapter is to resolve the problem of other minds, I shall begin with an outline of the epistemological sceptic's approach to one's claim to know that others are conscious or have particular mental states. He argues that because all that I can observe or be acquainted with in other people is behaviour, unlike in my own case, my knowledge of their conscious states must be indirect. If my belief is that, for instance, another person is in pain, then this has to be an inference, or justified as such, so that a question arises as to the grounds for it, the answer to which will depend on the view one takes of its logical status. Since the sceptic aligns himself with common sense in rejecting the behaviouristic theory by which pain and pain behaviour are equated, he would say that they are distinct phenomena, inferentially and causally linked but logically independent, so that the inference which links them is factual and in need of justification.

Factual inference figured prominently in the previous study of Hume's problem, in connection with which it should be noted that the sceptic about other minds is not, and cannot be, a sceptic about induction – on the contrary, the thrust of his thesis just is that what we believe about other minds is at best only weakly supported when compared not only with direct awareness of our own conscious states but also with our knowledge of the physical world, including the behaviour and actions of its human occupants. I know a great deal, after all, about other people's overt behaviour, where this includes making predictions confirmable by direct observation. But when it comes to *explaining* what they do, by way of inference to the particular contents of their inner life, then nothing counts as my direct observation or awareness of their conscious states.

How, then, are beliefs about other minds to be justified? A possible answer is that at least I have direct access to my own inner life, and that this enables me to argue by analogy to that of other people. But now the sceptic claims that this appeal to analogy makes a very feeble argument, and he may go on to suggest that it lacks all credibility. The reason is, I think, an interesting one, and it proceeds from the fact that when I draw conclusions about, say, another person's pain I am not starting from scratch, as it were, as if I have never encountered another human being before. When I ascribe pain I do so from within a belief system in which this other person is already recognised as such, which is to say as being similar to myself and other human beings in having an inner life. I interpret his behaviour in

terms of a limited range of alternative explanations, which in this case reduce, broadly speaking, to his being in pain or pretending to be, both of which involve his being conscious. What lends itself to my construing him as a person, as if it were not all complicated enough already, is that I do not 'see' people as physical objects and infer to their being conscious; rather, I 'see' them as looking out at me from their faces, on which the play of expression seems directly linked to their inner selves, just as I 'see' physical objects as having spatial depth. For the sceptic, however, this impression of direct awareness is misleading, and all our beliefs about other people are inferential and depend on appearance, behaviour and circumstances.

To see the difference this makes with regard to the argument from analogy, we may note that within the system it is rational, or at least natural, to discover connections between the consciousness and behaviour of oneself and other people, at least if the question of ultimate justification is disregarded. It is this question, however, with which the sceptic is concerned, so that the argument from analogy is required not just to operate within the system but to underpin it. Thus, the sceptic really does require us to start from scratch and not presuppose that other people have an inner life. This would be no mean feat, for I have to imagine encountering an individual without having any preconceptions about him as being conscious, so that I do not initially recognise him as a person. For instance, instead of saying that I see him walking towards me and crying out in pain when felled by a tree, I now have to say that I see a figure approaching and that it resolves itself into a bipedal object. A branch of a tree crashes onto it and it falls over. A hole appears in its top section, which is similar to my head and face, and sounds are emitted, like my own when I am in pain; therefore, this object is likely to be in pain – as an epistemic basis for our whole system of inter-personal beliefs this is not very promising, especially in view of the fact, often overlooked, that I cannot observe myself as I do other people, and in particular that I cannot see my face.

Let me try to sum up the sceptic's position. He argues, or I do on his behalf, that our knowledge of others' conscious states belongs within an explanatory system in which the existence of other minds is presupposed. If I see a distant human figure walking towards me, I know already that it is a conscious person, for walking depends on perceiving and on sensations of balance. What is happening, very clearly, is that my own perception of this human figure is being slotted into a pre-existing explanatory system, this being true of any contact I have with other people. If that is the case, there is no general justification for the inferential step I take from others' behaviour to their underlying consciousness, given the

absence of direct confirmation and the fact that particular justifications always presuppose that other people have minds. I could start from scratch, without any preconceptions, only if the hypothesis of non-consciousness was brought into play, making three alternatives to be weighed up; and this could be done, according to the sceptic, only if I inferred strictly from what was observable, from appearance and movement, and if I referred to it in terms that were neural between the alternatives. But then, or so it would be said, it is very difficult to see how I could ever justify ascribing consciousness to another person, the argument from analogy lacking all credibility when detached from the system.

If this account of scepticism about other minds is correct, then it helps to clarify the role of the argument from analogy in the armoury of the anti-sceptic. On the one hand it is assumed that the sceptic forces us to retrench at the level of arguing from one's own case; on the other it is recognised that such an argument is ineffective when cut off from support within a system, where it enters into inferring complex interconnections between one's own and other people's inner states and outward behaviour. My own approach will be to deny that this retrenchment is necessary. In the meantime, let us consider other rejoinders to the sceptic, one of which is to adopt the behaviouristic outlook which he sides with common sense in rejecting. It will then be argued that although I ascribe conscious states to myself, my ascription of any particular one of them to others is to be equated with those references to their behaviour on which it is normally taken to be based. Since any inference involved is only to what is observable, not to hidden mental phenomena to which I have no direct access, this resolves the epistemological difficulty and disposes of the sceptic. There are, however, two drawbacks to this heroic attempt at a solution, the first of which is that I *do* ascribe consciousness to others as to myself; and, secondly, that behaviourism is tantamount to solipsism. The reason is that if as a behaviourist I am able to say that I am conscious, but with the caveat that I can refer only to other people's outward actions and appearance when I say that they are, so that the meaning of the psychological predicates I use depends on whether the subject is myself or someone else, then I cannot also claim, appealing to symmetry, that the same is true for them. If that is what I imply, that they can refer to their own conscious states but only to my behaviour, then either I contradict myself, both asserting and denying that I am able to credit them with consciousness, or what I say about their self-ascription of consciousness has itself to be given a behaviouristic interpretation, where this is something that I fail to make clear. Put another way, the behaviourist is a solipsist in disguise, though I shall later explain, as already alluded to, that conversion to solipsism would remove a major

source of personal anxiety in everyday life, and is therefore to be recommended to devotees of the philosophy of pragmatism.

Since non-pragmatists still await a solution to the problem of other minds, we shall continue to examine the various candidates, using as our starting point one of the assumptions underlying behaviourism. The behaviourist assumes, in contradistinction to the epistemological sceptic, that if mental concepts are fully grasped when mental states are correctly self-ascribed, then the corresponding predicates cannot have the same meaning when attached to a subject other than oneself. One answer to this, which either rejects what the behaviourist assumes or attempts to circumvent it, is to argue that there is more to the essence of such concepts than the ability to apply them in one's own case. That, at least, is the view advanced by P.F.Strawson in his book *Individuals*, which we shall now consider in relation to his anti-sceptical thesis.

Strawson's attempt at a solution

The gist of that thesis is that the notion of a person is primitive, and that the ability to ascribe mental states to others is a pre-condition of being able to ascribe them to oneself. This is in conjunction with the view that psychological terms have the same meaning whether applied to oneself or to other people, so that behaviourism is to be rejected. Also rejected is the argument from analogy, for if I have to ascribe experiences to others in order to ascribe them to myself, then the argument from my own case presupposes other minds. A central claim is that the behaviour of others constitutes a logically adequate criterion for the occurrence of the inner experiences with which it is epistemically associated. This is held to be true not only for those predicates which straddle experience and behaviour but also for those which pick out only the experience, as with pain and other sensations, the associated external signs coming under a different description

If the notion of logically adequate criteria can be made respectable, then it would seem that it has the power to deprive the sceptic about other minds of any useful employment. So the following argument, if correct, would make him redundant: mental states may be self-ascribed only if they may be other-ascribed, the ascription criteria being logically adequate. Thus, the sceptic's doubts are idle because he cannot formulate them without using concepts that rely on logically adequate criteria for the ascription of mental states to other people. It would be as if he denied that we have any reason to believe that the colour red exists; and then it could be argued that he understands the word 'red' only if he accepts criteria for its application, from which it follows that his denial is idle, since he must believe

that red exists in order for his denial of it to make sense. This may seem to be a striking analogy until we remember that criteria for colour ascription take the form of seeing the colours themselves, whereas the criterion for another's pain is pain behaviour, which is not even similar to pain let alone identical with it.

It is with this distinction in mind that Ayer casts doubt on the cogency of Strawson's thesis, by citing the possibility of identifying someone's behaviour as being characteristic of the expression of pain, and yet denying that pain is what he feels. The behaviour and the sensation are logically separate, he says, and this would seem to count against the claim that criteria for the ascription of pain to others are logically adequate. What he does acknowledge, however, is that it may be worth considering what he takes to be Wittgenstein's view, to the effect that:

Though the liaison between the characteristic outward expression of an inner state and the inner state in question may fail in any particular instance, it is not logically possible that it should fail in all instances, or even in any high proportion of them. (1973, p.100)

That is what Ayer says, but it may be possible to add some detail to the view he outlines if we consider it in the context of a system. If someone gives every sign of being in pain, but if I do not accept that he is, then it has to be, or so it will be said, that in normal circumstances I would accept it, so that the present case must be exceptional in some way. Perhaps I have reason to believe that he is either pretending or in the grip of what used to be called hysteria, in which case I am not in breach of the rules of evidence for judging whether another person is in pain. But if, on the other hand, I discount these latter possibilities yet still insist that there is no reason to believe that he is in pain, so that I do not count the blood and the screams as evidence, though I acknowledge that normally one does, then it may be suspected that my grasp of the concept of pain is deficient in essential respects, so that my self-ascription of pain is compromised.

It is in this way, it seems, that it is thought that the notion of logically adequate criteria may be put to anti-sceptical use; but in any case Ayer now goes on to reject the claims that are made for it, and to prove his point he conducts a thought experiment to show that even if behavioural criteria are accepted, they are not logically adequate. He asks us to imagine a child brought up by machines which teach him to speak and induce in him the belief that they are conscious and have mental states that are similar to his, the point being that the child is systematically mistaken in his ascription of consciousness, or particular forms of it, to these machines.

Perhaps it is now time for me to join the fray, and to side with Ayer in his objections to Strawson's anti-sceptical thesis. His position is that the notion of logically adequate criteria of pain implies that a relation between statements obtains which is not deductive, since ascribing pain to others is not to be equated with a description of its physical manifestations; and yet, it is not inductive, either, so that it hovers between entailment and factual inference. The obscurity and uncertain status of this relation represents a weakness in Strawson's theory, he says; and it might be thought that another is given by the ease with which apparent counter-examples may be cited, including the one devised by Ayer himself. If Strawson is correct, after all, then solipsism is not a coherent view; and yet, I seem able to imagine ceasing to believe that other minds exist, with my references to my own mental states being unaffected.

What I now propose is to try to reinforce these objections by turning Strawson's thesis around and examining its converse: that self-ascription is a necessary condition of other-ascription. At first sight it may seem to be in line with common sense, as also with empiricism, to believe that I cannot other-ascribe pain unless I am able both to feel it and to identify it, as opposed to not being able to feel it or not being able to identify it.[4] And this, as far as I know, is correct as a factual thesis about acquiring the concept of pain, which involves forging semantic and cognitive links between one's own and others' pain and pain behaviour. As regards the details of this process, no doubt a child psychologist would be able to enlighten us, as also about the learning difficulties faced by a sufferer of congenital analgesia, or, with regard to learning the language of colours, by a child who has been blind from birth. Our concern, however, is with the logic of concepts, about which there seems to be an assumption carried over from the factual context of learning: that we can ascribe pain to others only if we have experienced it ourselves. This, however, is far too vague as it stands, as also with its commonly being said that some experiences are so special that we need to have them in order to know what they are like. One example would be that of giving birth, which we are said to understand only if we ourselves have had babies, it not being an adequate substitute that mothers, however articulate, should describe their experiences to us. If sentiments of this kind are to be elevated to a philosophical level, then perhaps in the present case it would have to be said that one's experience of pain is what gives mentalistic content to ascribing it to others, so that one's reference may go beyond the outward signs to the other person's pain.

---

[4] There actually is a rare genetic disorder, congenital analgesia, characterised by inability to feel pain.

Set against this, however, are such facts as that one does not need to feel pain simultaneously with ascribing it to someone else, nor is one required to have an exactly similar pain, for it is the case, surely, that I need not have poor dental hygiene in order to refer to another person's toothache. The indication here is that if my own experience of pain expands my cognitive horizons, thus enabling other people's pain to come within the orbit of my understanding, then this is a contingent fact about concept acquisition, not a necessary fact about the logic of pain ascription. Therefore, it is conceivable that without being able to feel pain I should be able to ascribe it to others. I am required only to be *able* to ascribe it to myself, in other words to be suitably disposed with respect to the self-ascription of pain, the difference between myself and others being that I never have occasion to exercise such dispositions. Thus, if I were to feel pain I would be able to identify it as such, my use of 'pain' also being correct in the sense that I do not refer to the wrong sensations as instances of pain. If to this it is objected that so constituted I cannot *know* that I am suitably disposed, then even if this is granted as a contingent fact, it still is the case that conceivably I do not doubt what my dispositions are and I claim to know that I would be able to refer to and describe my pains if the occasion were to arise.

Perhaps it may now be in order to clarify certain points. In examining the link between my own experience of pain and my coming to understand other people's, I have kept in mind the irreducibility thesis advanced in a previous chapter. As pointed out earlier, it is within an explanatory system that I make connections between the pain and pain behaviour of myself and others, and I am now suggesting that it is conceivable that I could ascribe pain to others without being able to feel it myself. But this is not to imply a resuscitation of the possibility that the meaning of such statements as 'he is in pain' is behaviouristically reducible. If I manifest understanding of what it is for others to be in pain, whatever the means by which I have acquired it, then my ascription of pain to them is mentalistic and refers not to their behaviour but to their experience of pain.

The upshot of these arguments is that although I do not need to feel pain in order to know that others do, I do need to be able to self-ascribe it, which is to say to identify it in myself should I feel it. Nowhere in this account is there any indication that in order to self-ascribe pain I need to accept behavioural and other external criteria for its other-ascription, even if it is granted that I need to be able to recognise them as such. The conclusion to be drawn, then, is that Strawson's thesis is incorrect.

An astounding fact about other minds

Since our advance against the sceptic has stalled, at least for the moment, perhaps it is time to take stock and then to re-group. The one thing we can agree with Strawson about is that if a psychological predicate has a mentalistic meaning when the subject is oneself, then this remains the case when the subject is someone else. Set against this, in some ways, is the fact that in our dealings with others our concern tends to be with their behaviour, albeit within an explanatory system in which we interpret it as that of persons, with all that this implies about their being conscious. Within that system it still is the case that sometimes we might as well be interacting with an automaton for all the interest that we take in the conscious phenomena of another person's mind. At the other extreme, of course, we may be acutely aware of, or at least very concerned with, another individual's particular experience, or even with what it is like for her to have it, where this is to manifest or strive towards an empathy we never feel with bats. That said, there clearly is a practical, everyday level at which we focus on another person's behaviour and actions, and with a view to their effect on us, or to predicting and controlling them, to which end even our awareness of that person's inner life may be subordinated.

This focus on behaviour throws light on the non-mentalistic semantic component of many of the terms by which everyday human psychology is described and explained. We say not only that a person is in pain but also that she is friendly, which refers to her demeanour, or that she is becoming angry, perhaps at having to be friendly while in pain, and here the reference is to emotion linked to behaviour, the particular feeling and action having to fall within a limited range of such phenomena, that which covers what it is to be angry. The fact is, however, that none of this helps to neutralise or limit the threat posed by the sceptic about other minds, for even when we refer directly to behaviour, as with friendliness, we imply that the subject is conscious, so that a question still arises as to how we know that she is.[5]

Continuing in this vein, let me now try to render as vivid and striking as I possibly can the reality of other minds, so that we may realise how much is at stake when the sceptic questions their existence. Suppose, then, that another person and myself are looking at a uniformly blue wall, and that our colour experiences are as similar as it is possible for such experiences to be. Then it is my contention that they may be *exactly* the same, and in the same way as my own such experiences from one moment to the next, so that they are qualitatively identical. To bring out the essence of what this means, let us consider objections

---

[5] It is for this reason that Ryle's *The Concept of Mind* leaves untouched the threat posed by the other minds sceptic.

to it, one of which concerns personal identity. The argument is that even if my perceptual experience is the same as the other person's, what this could mean is constrained by the fact that we are separate individuals, so that for each of us our unique self runs through our own experience like a name through a stick of rock. To this it may be replied, following Hume, that I am not acquainted with my self as a subject of whom experiences are predicated, for the self is not an experiential object. One could argue, too, that even if it *were* such an object, the experience of it could be both the same for all of us *and* distinguishable from other kinds of experience, for instance that of blue. There is, of course, the grammatical subject, as when one says 'I am immersed in blueness', but again this reference to myself may be exactly the same phenomenally as the other person's reference to himself when he uses the same sentence. That leaves dispositional differences, but a disposition, as already pointed out, is not an object of direct acquaintance. Again, then, we are back with my experience of blue being qualitatively identical with the other person's. It follows that the felt reality of that experience is *the same for both of us*.

This is a startling result, hence the italics, and there is more to come. I have focussed on qualitatively identical experiences among different individuals, but only in order to elucidate the point that the felt reality for each individual is the same. That point is wider-ranging, as may be shown if we now suppose not that my visual experience and another person's are exactly alike but that we are each looking at a different uniformly coloured wall, perhaps in separate rooms, so that we may be seeing different colours. Even as individuals we do, of course, see different colours, either at different times or present in the same visual field; and there is a sense in which the felt reality of visual perception remains constant through colour change, so that we may claim that it is the same in one person as in another who sees a different colour. By extension, it may then be argued that all my experiences partake of the same reality, which is identical from one person to the next. Thus, if the other person and myself now meet in the same room and face each other, then there are ways in which I may venture a description of my experience that he will recognise as true of his. For instance: that each moment of it is charged with realness; that it is *as if* I inhabit my head, a separate presence, and look through twin windows at my companion's eyes on me, the light reflected back as if from stained glass. This person is qualitatively identical with me as a subject of experience, the reality of which is the same for me as for him as we look at each other's different faces. There are *two* worlds, connected in the moment but entirely separate, and yet partaking of the same reality.

The point I wish to capture in these words is, as shown by my straining after it, very elusive and hard to pin down, which is why I have allowed myself some latitude in certain areas. If we are to fully grasp the point, then I think that we must now return to those cases in which different people have the same experiences. Suppose again that myself and another are seeing exactly the same when looking at a blue wall; then there is nothing in the visual experience itself by which it could be assigned an owner, and if my experiences at different moments were interchanged, either with one another or with the other person's, then this would not be discoverable from the experience itself. It is in making such claims that I permit myself some latitude, for the notion of an interpersonal exchange of experience has only a tenuous link with what is conceivable; but it does help to show that other people's conscious states are real in the way that mine are. This is obvious as a stated fact, as may be seen from trying to deny it, but it is one that, moments of empathy aside, we rarely focus on in any depth, or not as we would a familiar object seen under a microscope, the magnification revealing it as something strange and new. What should finally be said, allowing as bright a light as possible to strike the object, is that, amazingly, another person's every conscious moment, in all its immediacy and aliveness and in its quality of being real, *is exactly the same as mine*.

This is, as already indicated, an absolutely astonishing fact; and what it shows is that there is not just a gap but a chasm between premiss and conclusion when I infer to another's conscious state, the existence of which is not demonstrable, the same being true of the blanket belief, as it were, that consciousness resides in this other person. This is not to deny that we take for granted that other people are conscious, at least before we are confronted by the sceptic, but what it indicates is that the truth of solipsism is logically possible. This is, at the level of everyday belief, to pay lip service to bare possibility, and it would take some extraordinary event, perhaps the stuff of science fiction or nightmare, for me to be induced, while remaining sane, to really doubt that other minds exist. In answering the sceptic, then, what I need to demonstrate is not that the consciousness of others is entailed by the evidence but that it cannot reasonably be doubted, so that the solipsist in me is buried under the weight of such evidence and silenced by the particular judgements I make.

Epistemic perception of others

How, then, is this to be shown? In *The Possibility of Knowledge* Quassim Cassam advocates what he refers to as the perceptual model of other minds, whereby emphasis is placed on direct epistemic perception of other people as being in a particular mental state, as when they

are angry, in pain, and so on, as opposed to consciously inferring that they are. Similarly, we see physical objects as being three-dimensional. It is clearly the case, however, that direct observation has its limits, as when we rely on others' self-ascription of sensations, intention states, and so on, and Cassam does not deny that much of what we know of others is indirect and inferential. He says, 'it seems that neither inductive inference nor inference to the best explanation can put me in a position to *know* that others think or feel anything; the most I can conclude on the basis of either form of inference is that it's *probable* that there are other subjects of thought and experience'. (p.157) Wishing to disagree, I would argue that in opposing the sceptic we need not seek to demonstrate the existence of other minds but only to defend our conviction that they exist, this being something we take for granted as part of the structure of the explanatory system by which we judge that other people have one particular mental state rather than another, the possibility that they are not conscious subjects, so that the rock is hollow, not being considered.

Not only that but I also think that Cassam distinguishes too sharply between perceptual and inferential knowledge, if only for the reason that much of what I seem to directly perceive is a result of unconscious learning or was initially worked out by conscious inference, the scaffolding then being removed. Thus, the distinction is not ultimate, and if understanding informs perception, the two fusing together in perceptual experience, then that experience stands as proxy for what would otherwise be a process of inference. If, as adults, we encounter an unfamiliar object and move around it in order to work out its size and shape, then from the first moment we 'see' it as being three-dimensional, as we do its parts as they come into view, so that when we return to our starting point we now 'see' the object differently, that difference corresponding to what we have learned of the object's size and shape. That, at any rate, is the philosophical model of perception to which I subscribe, and it applies to our observations of other people, for on first encounter with another person we 'see' her as being conscious and, as Cassam would say, we see that she is in a particular mood, or at least we cannot help reading something into her expression and demeanour, and as we move around her, as it were, we build up a picture of her to which all our observations of her contribute and by which they themselves deepen in colour. The point here is that when the picture is thereby drawn and developed it is within an explanatory system in which connections are made between our own and others' mental states, behaviour and actions, such that when I see that another person is angry, which Cassam would refer to as epistemic perception, what I 'see' goes beyond what I see, so that my knowledge of other minds is *always* inferential.

That being the case, we cannot hope to refute the sceptic by appeal to epistemic perception; but Cassam would not agree, and he now goes on to consider a particular sceptical argument, one with which we are familiar in the form of the so-called paradox of the brain in a vat. Suppose, he says, that we can see that a college bursar is angry. If he is angry, then he is not a zombie, so do we imply that he is not? If so, then how do we know that he is not, given that his behaviour would be the same if he were? Cassam suggests that if I see that the bursar is angry it must be true that he is not a zombie, but I need not know that he is not, independently of seeing that he is angry. He admits that this is not a satisfactory response to the sceptic, who could simply insist that if I see that the bursar is angry, then I must know that he is not a zombie. The correct approach, or so I have tried to show, is to appeal to the notion of a system and deny that the sceptic can coherently apply the method of constructing hypotheses at variance with the claims to knowledge he wishes to undermine. It would mean that if he embarked on an epistemic voyage he could never set foot on the terra firma of solid belief but would almost immediately capsize into a sea of conflicting possibilities. But also, if believing and knowing belong within a system, so that belief is not anchored to evidence in the way that foundationalists think it is, then it should always be possible to scupper any sceptical hypothesis, such as that the bursar is a zombie. Quite simply, the evidence is overwhelming that zombies do not exist, and that is all there is to it.

Now that we are waving goodbye to Cassam, is there anything of his that we should take on board in these last stages of progressing towards a solution of the other minds problem? His approach is to take up a common-sense position and then to defend it against the sceptic, and although his defence of the proposition that other minds exist is very different from ours, we can learn from it and invite the sceptic to argue for his undermining of common sense, not by using the discredited method of sceptical hypothesis but by advancing arguments purporting to reveal defects in our reasoning about others' mental states. A special difficulty the sceptic faces, for there would be a difficulty anyway, is that he cannot also be a sceptic about induction and our knowledge of the external world. He needs to exploit the fact that we have no direct access to other minds for purposes of evidence and confirmation, and it has to be done in such a way as not to compromise his acceptance of these other forms of empirical knowledge, in particular our knowledge of other people's appearance and behaviour, or that of automatons if a contrast is to be made.

A countryside stroll with a robot

In investigating these matters, what I suggest is that we conduct a thought-experiment, our virtual laboratory being a future world in which humanoid automatons roam the countryside and help with conservation work, such as reclaiming lawns buried under concrete drives or restoring fields that were traditionally given over to monoculture, where they are employed in spraying chemicals that would be dangerous to humans. This is a world, then, in which the owners of automatons hire them out to nature reserve wardens or to suburban homeowners in receipt of a lawn reclamation grant. Projecting myself into this future time, I imagine that I am going for a countryside stroll along an elevated walkway above moving traffic belts and that I descry two distant figures approaching, one ahead of the other, the nearer of which I recognise as an automaton, judging by its stiff, mechanical gait and head swivelling. Already, then, my identification of this approaching figure as a programmed humanoid goes far beyond my limited perception of it in terms of shape, colour and movement, on the basis of which I expect it to resolve itself, as its details become visible, into a state-of-the-art robot. If I am asked to justify this expectation in the ordinary way, then I may do so by appealing vaguely to the broad sweep of the relevant knowledge I have; but also, and more specifically, I may adduce the fact that very often in the past I have seen distant moving objects that are distinctive in certain ways resolve themselves into automatons. Even in the absence of such experience, I may appeal to what I have learnt from testimony; hence my confidence in this particular case.

What should now be noted is that the leap I make from the sight of this tiny, distant figure to the belief that it is an approaching automaton is one that can only be justified within an explanatory system. If a sceptic about induction, or about knowledge of the external world, insists that the grounds for my belief should not go beyond the sighting itself, neutrally described to avoid any charge of question-begging, then it is not just that I cannot comply but that nothing would count as meeting his demand. It is necessarily the case, as already argued for in so many ways, that my particular empirical beliefs belong within a system in which the legitimacy of such beliefs in general is presupposed. The point is that the sceptic about other minds has to acquiesce in this thesis, given that he cannot also be a sceptic about induction, who in any case would refute himself. It is worth noting that the thesis owes nothing to coherence theory, for it relies on the notion of fundamental conditions of discourse and of knowledge of the world, such that empirical beliefs are self-sustaining within a system. Thus, in the present case it may be said that I just *have* the belief about the

automaton, and if I justify it by appeal to experience then again I just *have* the beliefs which enter into that justification.

Still in this future world, let us now suppose that after I draw level with the automaton, which gives me a wink without breaking its robotic stride, I focus on the other distant figure, which I recognise as human from the fact that it is running, something which automatons are not designed to do. Again, then, it is clear that what I see, limited as it is, induces in me beliefs and expectations which I can justify by appeal to my experience of people, in particular what they look like from afar, together with what I know of the difficulties faced by owners of automatons in keeping their charges under control. For the person drawing near is the robot's owner, as I have already guessed and as she confirms when she tells me, cursing and fuming, what she intends to do to it when she catches it. She is gasping for breath and holding her side, and she explains that she is waiting for the pain of a stitch to ease off. In understanding and assessing what she says, in the light of her behaviour in the given circumstances, I must interpret it from within an explanatory system in which the existence of other minds is presupposed, so that all that can be in question is the particular belief it is rational to hold, which is this case is that she is genuinely in pain. At no point does it occur to me to argue from my own case, perhaps with the help of running on the spot until I get a stitch.

Direct assault on scepticism

If we now consider the insights to be gained from this excursion into the future, then it is clear that the sceptic about other minds is not entitled to demand that we start from scratch when attempting to meet his challenge, and this is to say that we misread the nature of the problem if we resort to the argument from analogy in making that attempt. It would be premature, however, to start celebrating our victory over him at this early stage in the campaign, for it is now open to him to withdraw to his basic position: that statements about other minds are fatally flawed in that, unlike those about the physical world, they are not directly verifiable, this being part of the logic of the concepts involved. In preparing our rejoinder to this, we need first of all to characterise the scepticism with which we are dealing as being radical in nature, so that the sceptic maintains that we have *no* reason to believe in other minds. Let us now mount a frontal assault on this position by asking why it is that direct verifiability is thought to make physical object statements secure. Since a physical object is three-dimensional, I cannot observe the whole of it at once, and in many cases all I ever see of it is its external surface. What is true is that I can see the whole of it, or the whole of the

outside of it, if I turn it in my hand, for instance, or walk around it, unless it is the moon, and in this way none of it need be hidden from me, whereas the mental states of other people are always hidden. This is not in dispute, but its epistemological import needs to be brought under scrutiny. If I believe that there is a corridor outside that door, I believe it now, despite the absence of any direct observational evidence. Such evidence may be obtained by the simple expedient of opening the door, but if this is taken to confirm my belief, then it does so only if I remember believing it. Also, when I open the door my stating of the belief in question is a past event, my present sighting of the corridor being confirmatory only on the assumption that it existed at the time I stated my belief. It is by way of such assumptions, as I earlier argued in detail, that induction is fundamental, this being implied in the claim that certain basic concepts are transtemporal. At such a basic level, indeed, it is arguable that this continuity assumption, as it may be called, is intrinsic to the concept of the present moment, in the context of believing in the present occurrence of an event, or existence of a state or property, even when it is an immediate object of perceptual acquaintance. That said, there may still be a great deal of force in the claim that *strictly* speaking such beliefs are not directly verifiable.

Still very much on the offensive, let us now exploit some of the previous arguments about the epistemic significance of memory. If the approach we have taken is correct, then the process of arriving at a belief and directly verifying it is wholly dependent on memory in one form or another. Thus, in order to verify that this is a computer monitor on the table in front of me, I could crawl around to examine the back of it, in which position I would lose sight of the front, so that nothing is confirmed unless I take my memory of the front to be veridical, in which case I could have saved myself the effort by relying on my memory of the back, if previously observed, or on testimony, or on knowing that this object would not behave as a monitor unless it was one. The point is that none of this would be possible without reliance on memory and going beyond what is given, the same being true in simpler cases such as the earlier one of expecting to see a yellow disc glowing in the dark. And yet, the past is no more directly accessible to me than the contents of another person's mind. Generalising from these examples, it is possible to argue that although there is a familiar sense in which physical object statements are directly verifiable in a way that those about other minds are not, this does not yield a distinction that the sceptic about other minds can legitimately exploit.

Now that the radical sceptic is in retreat, it is time to consider another comparison by which it is claimed that statements about other minds are shown to be defective: that between

such statements and those about the conscious phenomena of one's own inner life. Again, it is not in dispute that one's beliefs about the inner lives of others are evidentially indirect; but let us now turn to what passes for direct knowledge in one's own case, by way of the stock example of one's knowing that one is in pain. In an earlier chapter I maintained that authoritative self-ascription of pain is contingent upon one's understanding of the ascription, which involves memory and being suitably disposed, so that it goes beyond the immediately given. All this, together with an irreducible semantic or cognitive element, makes the difference between having a pain and knowing that one has it, if such knowledge is verbally expressed. Not only does knowing that one is in pain go beyond the sensation itself, but a great deal of one's self-knowledge is indirect in more obvious ways, even with regard to pain, as with one's belief that one felt pain yesterday, or a moment ago; and, too, there is all that one believes about one's intentions, one's moods and emotions, not to mention the difficulty of capturing a particular thought, which one can only do with a different thought.

Enough has been said to remind us, harking back to an earlier chapter, that the notion of incorrigibility, in the present case of authoritative self-ascription, should not go unexamined, and there is the fact, too, that much of what we believe about ourselves has no pretensions to being authoritative. We need not pursue any of this, given that we are treating of radical scepticism about other minds, for it follows not at all from self-knowledge, however analysed, that there is no reason to believe that other minds exist; and unless the sceptic can mount an effective challenge, then we are entitled to maintain that on the contrary there are compelling reasons.

Leaving the radical sceptic to consider his options, let us now ask whether a less extreme form of epistemological scepticism about other minds is worth considering, or has been considered. What is often discussed, after all, is whether the credentials of the argument from analogy can be established by those who make claims for it as a defence against the sceptic, or whether it is the weak argument it seems to be, rather than no argument at all; and the indication here is that there may be a milder form of scepticism, in which it is queried whether the evidence warrants the certainty with which we attribute particular mental phenomena to other people. But now, there are several difficulties here, apart from that arising from the obvious fact that very often we are *not* certain about other people. There is, for instance, the fact that the moderate sceptic, like his radical counterpart, is not entitled to insist that we start from scratch in grounding our beliefs about other minds. Rather than attempt to deny us access to the explanatory system in which such beliefs have purchase, he must show, if he can, that the inferences by which we arrive at them are less well-founded

than we think. If to that end he clings to the indirectness of the evidence on which such inferences are based, then he can undermine them only if it can be shown that our rebuttal of the radical sceptic, with regard to evidence being indirect, is not effective against himself. A difficulty he faces, which perhaps is insoluble, is that the explanatory system *already* accommodates the multiplicity of ways in which indirectness of evidence imposes epistemic constraints on the beliefs in question.

Suppose, for instance, that I am a research scientist doing laboratory work with a colleague, and that I form the hypothesis that hand contact with the flame of a lit Bunsen burner is painful. We now devise an experiment in which I place my hand in the flame, my colleague taking notes, and then she does the same, with me taking notes if I still can. If the experiment now begins and I scream and leap into the air, my colleague reacting similarly in her turn, then I directly confirm the hypothesis in my own case, but only indirectly in my colleague's. The difference this makes is that I know, with maximal certainty, that I felt pain and still do, whereas it may be, for instance, that I have reason to suspect that my colleague suffers from congenital analgesia and pretended to feel pain because she does not wish her condition to be publicly known. And there are, of course, many cases in which even more of a difference is made, for there are countless ways in which we are less certain about other people than we would be about ourselves. Given, then, that the system already makes allowances for evidential indirectness and uncertainty, as manifested in such differences, how is the moderate sceptic to proceed and what exactly does he wish to challenge? Would he claim that most people injudiciously compare others to themselves, for instance that there are many cases in which, unlike in the present one, we completely ignore the possibility of dissimulated congenital analgesia when someone gives every sign of being in pain? But now, such a possibility and the claim about ignoring it would seem to lie within the system, not outside it, for what is implied is that deviations from the norm are sometimes overlooked, this being a proposition which presupposes the existence of other minds and the legitimacy of the evidence thereof.

Continuing in our quest to find an epistemological niche for moderate scepticism, let us now ask whether a sceptic about other minds could moderately insist that we take more seriously the possibility of solipsism. Since I have already argued that no amount of evidence that other minds exist entails that existence, I have to agree, indeed claim, that solipsism is a logical possibility; but this does not make me a sceptic, not even a moderate one, for logical possibility cannot of itself sustain any degree of doubt. It is possible, after all, with regard to the earlier laboratory experiment, that my fellow scientist not only suffers from congenital

analgesia and pretended to feel pain but that she hypnotised me, unbeknown to myself, into believing that my hand was in the flame, which surreptitiously she had just switched off, and into reacting as if in pain, which I now believe that I was, her motive being to spare me suffering and at the same time indulge my obsession with treating as a hypothesis a glaringly obvious fact. Since, however, the case is such that this possibility lacks any evidential support, I cannot coherently entertain it, and the reason is that if evidence is not required then *anything* is possible, including that the experiment never took place and that my colleague never existed. Logical possibility on its own has no epistemic significance and none in epistemology either.

If this is correct, then the sceptic is required to give reasons why the possibility of solipsism should be taken more seriously, and again we have to ask whether in the context of a system these would be internal, in the sense that they would counter the evidence that other minds exist, the strength of which would be acknowledged. One may imagine, if the sceptic is myself, that I come to believe that it has been revealed to me in a vision that I am the only conscious human being, the others all being automatons programmed by extraterrestrials, no expense spared, to keep me company. Such scepticism, however, would be neither moderate nor philosophical, so we need concern ourselves with it no further. There has been, I think, only one chance for the moderate sceptic to prove that he exists, and to take it he had to show that our strictures against the radical sceptic do not apply to himself, so that he is justified in insisting that we take seriously, on the basis of evidential indirectness, the possibility of solipsism. One might ask how seriously we are meant to take it, but since I do not think that we should take it seriously at all, I conclude that there is every reason to believe that other people exist, the one exception being the moderate epistemological sceptic about other minds.

Perhaps we should now ask in these closing paragraphs whether the preoccupation on the part of the sceptic and his opponents with knowledge of other minds being indirect can be accounted for, given that evidential indirectness is not problematic in the way that they take it to be. There are, I think, a number of factors, one of which is that it is assumed to be a consequence of the unthinkability of solipsism that it must be logically impossible, which it cannot be if the evidence for other minds is indirect. What is overlooked here, if an earlier claim is correct, is that many of our beliefs about the world are held with as much certainty as we attach to deductive conclusions, for instance our belief that other people are conscious when they seem to be, even if we may be less sure about the particular form that their consciousness takes. Secondly, it is possible that epistemological and semantic issues are sometimes confused, associated with which is that the means of acquisition of certain

concepts is taken to determine or constrain their semantic content. Thus, it is commonly said that one's grasp of psychological concepts depends on one's experience of applying them in one's own case, as with the connection between feeling and understanding pain. A query is then raised as to whether we can really understand what it is for another person to be in pain, given that our grasp of the concept is conditional upon our feeling pain ourselves, and in the context of our evidence for other people's being necessarily indirect, since it is only our own pain that we can feel. There is a misconception here which appeal to the irreducibility thesis may help to dispel, for it is only a contingent fact that I need to feel pain in order to understand the concept. Thus, there is nothing problematic, given that I have that understanding, however acquired, about my being able to ascribe pain to others on the basis of their behaviour and external circumstances. Finally, another contributory factor consists in the fact that there is taken to be an important distinction between direct and indirect evidence, whereas I have maintained, as part of the structure of this thesis, that such a distinction is not fundamental, unlike the fact that all propositions go beyond what is given.

The mystery remains

Perhaps we may now permit ourselves, having met the sceptic's challenge, to relax the arguments and give recognition to the elemental intellectual forces by which epistemological dynamics upthrust us into philosophy. These forces, after all, do not become quiescent just because the sceptic, who inhabits each of us, has been successfully engaged with. There are always fault lines in our awareness of the world, with all its discontinuities, and not even the seismic shift of resolving Hume's problem or that of other minds can still the process by which philosophical stresses build up; and if it is a permanent resolution one seeks, one might as well expect the Earth's tectonic plates to grind to a halt. There is a reason for this, to do with the nature of the world and its hidden structures, of which the surface appearance takes the form of the unanswerability of certain questions over aeons of philosophical time. How can an event now occurring, the snap of a camera shutter or a moment in the birth of a baby or a star, fall away into nothingness in the very act of playing itself out? How can I have so much power that entire worlds are at the mercy of my every perceptual whim, as when I gaze at the moon and then look away? And yet, how can I have so little that each moment dictates its own details, that always I blankly await the next, or turn my face into the shoulder of the familiar, and that my death will make no difference?

And then, to cap it all, there are other people. How can I be so *unreconciled* in my encounters with them, caught between staring at bodies and looking at fellow human beings,

except that their bodies are all that I can see? It is not as if I have any realistic prospect of resolving them into purely physical objects, to be approached with extreme caution but free from any feeling of awkwardness. Once I quit my study, as Hume would say, I abandon all hope of solipsism, other people being too insistently *there*, pressed close against the windows of their eyes, for me to risk taking them to be automatons, in case I start treating them as such. Immortality is, I dare say, a consolation of solipsism, albeit short-lived, but the reality is that other people are always with us, this having its own consolations, not least that we can always seek them out if we need to be alone together, rather than scan the sunless void at the edge of an empty world. Other people are, truth to tell, necessary, as is easily shown if we turn again to the teddy bear, that indispensable companion to the philosopher of mind. For the wisdom that this cuddly toy imparts as we soothe it and scan its face for understanding and communion, turning it this way and that in the morning light, is that if there were not other people we should have to invent them; and that this in any case is what we do, just as, even before it rises into consciousness, we hallucinate the Sun.

Bibliography

Ayer, A.J. 1959. *Philosophical Essays* (London: Macmillan)
—— 1972. *Probability and Evidence* (London and Basingstoke: Macmillan)
—— 1973a. *The Central Questions of Philosophy* (London: Weidenfield and Nicolson)
—— 1973. *The Concept of a Person* (London and Basingstoke: Macmillan)
—— 1985. *Wittgenstein* (London: Weidenfield and Nicolson)
Blackburn, Simon. 1973. *Reason and Prediction* (London: Cambridge University Press)
Baker, G. P., and P.M.S.Hacker. 1984. *Scepticism, Rules and Language* (Oxford: Basil Blackwell)
—— 1985. *Wittgenstein: Rules, Grammar and Necessity* (Oxford: Basil Blackwell)
Cassam, Quassim. 2007. *The Possibility of Knowledge* (Oxford: Oxford University Press)
DeRose, Keith. 1999. Solving the Sceptical Problem. In: DeRose, Keith; Warfield, Ted A. eds. *Skepticism: A Contemporary Reader* (Oxford: Oxford University Press)
Dretske, Fred. 1999. Epistemic Operators. In: DeRose, Keith; Warfield, Ted A. eds. *Skepticism: A Contemporary Reader* (Oxford: Oxford University Press)
Fisher, R. A. 1947. *The Design of Experiments* (Edinburgh: Oliver& Boyd)
——1956. *Statistical Methods and Scientific Inference* (Edinburgh: Oliver& Boyd)
Goodman, Nelson. 1983. *Fact, Fiction and Forecast* (Cambridge, Mass: Harvard University Press)
Grayling, A.C. 2008. *Scepticism and the Possibility of Knowledge* (London; New York: Continuum)
Grice, H.P. 1957. 'Meaning', *Philosophical Review*
Hattiangadi, Anandi. 2007. *Oughts and Thoughts* (Oxford; Oxford University Press)
Howson, Colin. 2000. *Hume's Problem* (Oxford: Clarendon Press)
Hume, David. 1975. *An Enquiry concerning Human Understanding*, ed. Nidditch (Oxford: Oxford University Press)
—— 2003. *A Treatise of Human Nature*, edited and abridged by John P. Wright, Robert Stecker and Gary Fuller (London: Everyman)
Keynes, J. M. 1973. *A Treatise On Probability* (London and Basingstoke: Macmillan)
Kripke, Saul A. 1982. *Wittgenstein on Rules and Private Language* (Oxford: Basil Blackwell)
Lange, Mark. 2002. 'Okasha On Inductive Scepticism', *The Philosophical Quarterly*, Vol. 52. No. 207
Law, Stephen. 2004. 'Five Private Language Arguments', *International Journal of Philosophical Studies* 12, no.2
Martin, C.B. and Heil, J. 1998.'Rules and Powers', *Philosophical Perspectives,* Vol. 12 *Language, Mind and Ontology*, 283-312
McDowell, John. 1998. *Mind, Value and Reality* (Cambridge, Mass.; Harvard University Press)
—— 1998. Response to Crispin Wright. In: Smith, B.C, Wright, C and Macdonald, C. eds. *Knowing Our Own Minds* (Oxford; Oxford University Press)
McGinn, Colin. 1987. *Wittgenstein on Meaning* (Oxford: Basil Blackwell)
McGinn, Marie. (2007). *Routledge Philosophy Guidebook to Wittgenstein and the Philosophical Investigations* (United Kingdom: Taylor and Francis Ltd)
Miller, Alexander. 1998. *Philosophy of language* (London: UCL Press)
Moore, G.E. 1959, *Philosophical papers* (New York: Collier Books)
Nagel, Thomas. 1979. *Mortal Questions* (Cambridge: Cambridge University Press)
Okasha, Samir. 2001. 'What Did Hume Really Show About Induction?' *The Philosophical Quarterly*, Vol. 51. No. 204.

Pears, David. 1988. *The False Prison: A Study of the Development of Wittgenstein's Philosophy* (Oxford: Oxford University Press)

Quine, W.V.O. 1953. *From a Logical Point of View* (Cambridge, Mass.: Harvard University Press)

Stern, Robert. 2000. Transcendental Arguments and Scepticism (New York: Oxford University Press)

Stove, D. C. 1986. *The Rationality of Induction* (Oxford: Clarendon Press)

Strawson, P.F. 1952. *Individuals: An Essay in Descriptive Metaphysics* (London: Methuen)

Stroud, Barry. 2000. *Meaning, Understanding, and Practice* (Oxford: Oxford University Press)

Weatherford, Roy. 1982. *Philosophical Foundations of Probability Theory* (London, Boston, Melbourne and Henley: Routledge and Kegan Paul)

Williams, D.C. 1947. *The Ground of Induction* (Cambridge, Mass.: Harvard University Press)

Williams, Michael. 1996. *Unnatural Doubts: Epistemological Realism and the Basis of Scepticism* (Princeton, New Jersey: Princeton University Press)

Wittgenstein, Ludwig. 1969. *On Certainty* (Oxford: Basil Blackwell)

Wittgenstein, Ludwig. 1976. *Philosophical Investigations* (Oxford: Basil Blackwell)

Wright, Crispin. 1984. 'Kripke's Account of the Argument Against Private Language', *The Journal of Philosophy* 81: 759-77

—— 1998. Self-Knowledge: The Wittgensteinian Legacy. In: Smith, B.C; Wright, C; Macdonald, C. eds. *Knowing Our Own Minds* (Oxford; Oxford University Press)

—— 2002. 'Critical Notice of Colin McGinn's *Wittgenstein on Meaning*', in *Rule-Following and Meaning*, ed. Miller, Alexander, and Crispin Wright (Chesham: Acumen)