Laudan, L. [1980]: 'Why Abandon the Logic of Discovery?', in T. Nickles (*ed.*), *Scientific Discovery, Logic, and Rationality*, Dordrecht: Reidel.

Laudan, L. [1996]: *Beyond Positivism and Relativism*, Boulder: Westview Press.

Lyell, C. [1833]: *The Principles of Geology*, 1st edn, London: John Murray.

Martin, E. and Osherson, D. [1998]: *Elements of Scientific Inquiry*, Cambridge: MIT Press.

Neyman, J. and Pearson, E. [1933]: 'On the Problem of the Most Efficient Tests of Statistical Hypotheses', *Philosophical Transactions of the Royal Society*, 231A, pp. 289–337.

Nozick, R. [1981]: *Philosophical Explanations*, Cambridge, MA: Harvard University Press.

Osherson, D., Stob, M. and Weinstein, S. [1986]: *Systems that Learn*, Cambridge, MA: MIT Press.

Putnam, H. [1963]: 'Degree of Confirmation and Inductive Logic', in A. Schilpp (*ed.*), *The Philosophy of Rudolph Carnap*, LaSalle: Open Court.

Putnam, H. [1965]: 'Trial and Error Predicates and a Solution to a Problem of Mostowski', *Journal of Symbolic Logic*, 30, pp. 49–57.

Quine, W. V. [1992]: *Pursuit of Truth*, Cambridge, MA: Harvard University Press.

Ruse, M. [1979]: *The Darwinian Revolution: Science Red in Tooth and Claw*, Chicago: University of Chicago Press.

Schulte, O. [2000]: 'Inferring Conservation Laws in Particle Physics: A Case Study in the Problem of Induction', *British Journal for the Philosophy of Science*, 51, pp. 701–806.

Spohn, W. [1988]: 'Ordinal Conditional Functions: A Dynamic Theory of Epistemic States', in B. Skyrms and W. Harper (*eds*), *Causation in Decision, Belief Change, and Statistics II*, Dordrecht: Kluwer.

Suppes, P. [1998]: 'Review of Kevin T. Kelly, *The Logic of Reliable Inquiry*', *British Journal for the Philosophy of Science*, 40, pp. 351–54.

Trudeau, R. [1987]: *The Non-Euclidean Revolution*, Boston: Birkhäuser.

Valdés-Peréz, R. and Erdmann, M. [1994]: 'Systematic Induction and Parsimony of Phenomenological Conservation Laws', *Computer Physics Communications*, 83, pp. 171–80.

van Fraassen, B. [1980]: *The Scientific Image*, Oxford: Clarendon Press.

2
The Theory of Knowledge: Saving Epistemology from the Epistemologists

Adam Morton

ABSTRACT

This is a very selective survey of developments in epistemology, concentrating on work from the past twenty years that is of interest to philosophers of science. The selection is organized around interesting connections between distinct themes. I first connect issues about skepticism to issues about the reliability of belief-acquiring processes. Next I connect discussions of the defeasibility of reasons for belief to accounts of the theory-independence of evidence. Then I connect doubts about Bayesian epistemology to issues about the content of perception. The last detailed connection is between considerations of the finiteness of cognition and epistemic virtues. To connect the connections I end by briefly discussing the pressure that consideration of social roles in the transmission of belief puts on the purposes of epistemology.

1 *Introduction*
2 *The shortest possible history*
3 *Knowledge: skepticism to reliability*
4 *From defeasible intuitions to the autonomy of evidence*
5 *Bayesianism to perceptual content*
6 *Finiteness to virtue*
7 *The structure of belief: from holism to foreground/background*
8 *Conclusion: from social networks to the biology of science*

1 Introduction

When philosophers of science address questions about confirmation, evidence, knowledge and rational belief they are usually addressing special cases of more general epistemological questions. But the flavor of the discussion is often quite different. Philosophers of science are concerned with a particular complex of belief acquiring institutions and practices. They want to know how they work, how well they work, and how they can be kept in good working order. Mainstream epistemologists, on the other hand, are usually motivated by concerns coming from within philosophy, or from pure intellectual puzzlement. This can make epistemology a very academic business. There are still

people out there wondering if they should believe in furniture. Yet epistemology has been flourishing recently; there's a creative atmosphere. The aim of this survey is to suggest developments from recent epistemology that ought to be of interest to general philosophers of science. My exposition will be biased towards bringing out connections between the developments. It ignores much of importance.

2 The shortest possible history

In 1950 the theory of knowledge was about to emerge from its pre-post-war form. Knowledge consisted in certain beliefs based on analytic truths and on sense data. Skeptical problems were unworrying because unsolvable. But things were about to change. Within fifteen years Quine had demolished the analytic synthetic distinction and the idea that to justify a belief is to base it on incorrigible evidence. Austin had fatally derided the idea of sense data. Gettier had pointed out the gap between knowledge and justified true belief. A new consensus reigned. Beliefs are justified if they cohere with the whole structure of a person's belief. Skepticism is based on a foolish obsession with certainty. Knowledge is belief that is true, justified, and has some exciting extra ingredient that needs to be discovered. The main tasks for epistemology are to find the ingredient, and to trawl through the history of philosophy explaining how the problems disappear when we abandon old hang-ups.

Then, some time in the 1980s, the consensus began to unravel. Quinean holism (now often called 'coherentism' to save it from sounding like a kind of alternative medicine) had been around long enough that it was noticeable that no one had given a clear explanation of its central terms. Skepticism found new forms. The magic ingredient for knowledge was never isolated. Most important of all, the hopeless search for the missing ingredient produced a deep contrast between internalist and externalist accounts. The former focus on inferential processes leading from one belief to another, which are subject to rational evaluation. The latter focus on causal links between facts and beliefs, which are evaluated in terms of their tendency to produce true beliefs. The contrast can be seen crudely with beliefs based on clairvoyance, or social intuition. Suppose a person is equipped with a faculty for telling what horses will win tomorrow's races or what people are dangerous. Suppose that the person has no reason to believe that the faculty is effective, and even has reasons for skepticism, but that in fact it really does deliver true beliefs. Then the resulting beliefs will be condemned as unjustified by internalists and praised as reliable by externalists.

So new issues began to appear. What would a non-rhetorical coherentism actually look like? How do we adjudicate or compromise between internalist

and externalist accounts? How do any of these points of view accommodate the post-Quinean realization that epistemology is one set of beliefs among others, characterizing the information-gathering capacities of one peculiar species? These have proved to be more fruitful and interesting questions than the ones they supplanted.

3 Knowledge: skepticism to reliability

Philosophers of science are often happy to talk loosely of knowledge in connection with any well-founded belief or theory. Sometimes it doesn't even have to be true. Not so epistemologists. Truth and justification are not enough; the best-founded belief may fail to be knowledge. And indeed this is in accordance with the way the everyday word 'knowledge' is used. So much the worse for the ordinary word, perhaps: it may seem not to matter for the scientific aim of getting true beliefs about nature, and of having an assessment of the relative faith we should have in different such beliefs. But in fact there is a point to holding out for more than good evidence. Consider, for example, a very academic-seeming skepticism which admits that we have good evidence for many of our beliefs, in fact that one would be a fool not to take this evidence as sufficient for accepting them, but insists on doubting that all this amounts to knowledge. Earlier skepticisms along these lines withheld the Knowledge label on the grounds of lack of certainty. No wonder philosophers of science ignored the issue. More recent writers have applied the same fundamental intuitions to more interesting claims. Two in particular. The first is contextualism, the idea that there are additional argument-places in the 'should believe' relation, so that given specific evidence belief in a proposition may still be in some sense relative to other factors in a person's situation. The simplest such contextualism makes the threshold for belief depend on the standards of the enquiry in which it is embedded. One can also make it depend on the conversational context in which the question of its truth arises. However the relativization is worked out, the effect will be that a person does not know something unless they have managed to rule out *some* knowledge-defeating possibilities. Which skeptical possibilities have to be ruled out depends on the context.

Why should we give a belief a different status in different contexts, though the available evidence is the same? To see the point we must move to the third-person perspective, and think in terms not of an agent evaluating her own beliefs but of a third-person assessment of the information provided by another. Suppose, for example, a theoretician absorbs some unwelcome results reported by an experimentalist and then passes the news on to another theoretician. Suppose that the same explicit evidence is available to all three. Yet there may be *potential* evidence that is relevant to their beliefs.

The experimentalist should have checked whether the equipment is malfunctioning in one way or another. And in fact she should have ruled out some extremely unlikely possibilities which, assuming they have not turned out to obtain, she will not even mention to the theoretician nor count in their common stock of evidence. (How often in *Physical Review Letters* do you read 'and by the way there were no mouse droppings on the leads'?) The first theoretician, for her part, should have checked whether the experimentalist was sober and speaking sincerely and that it was not April Fool's Day. But she need not worry about the mouse droppings, any more than the second theoretician need worry about the sobriety of the experimentalist. We can sum this up as follows. The first theoretician trusts what the experimentalist reports as long as she believes that the experimentalist knows that the result is as reported. In classifying the experimentalist as a knower she is assuming that the experimentalist has checked out various possibilities, most of which she is in no position even to describe. Similarly the second theoretician trusts what the first theoretician says as long as he makes a different set of assumptions about the kinds of things she has checked out. Knowledge is relative to context because the point of classifying someone as a knower demands different things depending on factors besides the explicit evidence available. (What factors, exactly? That is controversial. Suppose that the second theoretician mentions the possibility of mouse droppings on the leads to the first. Suppose that the first theoretician realizes that she has no idea whether this possibility was excluded. Some writers would argue that this change of merely conversational context is enough to make the first theoretician's belief cease to be knowledge. The crucial question, from the present perspective, is whether it has become less rational to rely on her testimony once this question has been raised.)

On the new approach to skepticism the classification of agents as suppliers of information is crucial. This emphasis links it to another recent tendency, which sees the vocabulary of epistemology, especially 'know' and its variants, as playing a social role in the evaluation, transmission, and storage of information which one person gains from the epistemic labors of others. It also links it to the *externalist* strand in recent epistemology, which emphasizes factors which link a person's beliefs to the facts, whether or not they are known to the person. Inference from evidence is just one among many such links, and a person's knowledge may be both greater and less than she would be rational to think it is. One reason for an external perspective is that we want not only to evaluate our own beliefs but also to assess the beliefs of others as sources of information. We use others as repositories of information and as information-gathering devices, and we want to know whether they are reliable. So there is a point to a conception of knowledge as belief that

results from a reliable process.[1] This conception is obviously in the same general spirit as a testimony-oriented conception, and is not necessarily incompatible with a contextualist one. All can be in a general way externalist, in that the classification of a person's grasp of a fact can depend on factors unknown to the person and outside her control.[2]

4 From defeasible intuitions to the autonomy of evidence

In older epistemological theories, and in formal theories of confirmation, the force of a reason for a belief *b* is absolute in the following sense: among all the information available to a person some counts as evidence relevant to *b*, and relative to this evidence there is a determinate degree to which the person is justified in believing *b*. In particular, if the evidence makes *b* certain, then no discoveries can alter the fact that relative to that evidence *b* is certain. Simple coherentist epistemologies make the principle unsustainable unless evidence is understood to include all of a person's prior beliefs. The concept of evidence and the category of a priori knowledge then become pretty pointless. (Some would be sorrier to lose one than the other.) But we can retain a modified form of the absoluteness of evidence while maintaining a generally coherentist attitude if we reformulate, using the idea of defeasible reasons, reasons which have force in the absence of specific countervailing considerations. Theorists of perception and of the a priori have both used this idea. Take mathematical proof, for example. Real proofs are never completely rigorous, and they have premises, which we usually cannot justify. Epistemologists sometimes write as if this approximates an ideal situation in which the premises are indubitable—just by considering them a person can establish their truth irreversibly. But we do not have to model the situation in this way. Instead, we can take the person's grasp of the premises as defeasibly certain: until and unless some consideration emerges that undermines the premises or a hole appears in the deduction the person can reasonably take the result as proven. In fact she can take it as *absolutely* certain, beyond a doubt. Thus

---

[1] There are hard questions about the definition of reliability. The contrast between knowledge-as-reliability and justification-as-reasons increases when one appreciates Goldman's point that justification needs a comparison of a wide range of actual or probable situations while reliability needs a comparison of a deep slice of possible alternatives to the exact actual situation. They're orthogonal.

[2] For a summary of the first phase of work on the definition of knowledge, see Shope ([1983]). Williams ([1977]) and Kirkham ([1984]) have doubts about its significance. Unger ([1975]) has haunted later writings on skepticism, and is a distant source of contextualism. More direct sources are Stein ([1976]) and Dretske ([1981a]). Recent contextualist theories are Williams ([1992]), deRose ([1995]), Lewis ([1996]). The whole contextualist approach is queried by Vogel ([1999]). For the third-person perspective on knowledge, see Welbourne ([1986]) and Craig ([1990]). For externalism see Goldman ([1986], [1999a]). For links between epistemic externalism, semantic externalism, and the third-person perspective, see Burge ([1979]).

eighteenth century natural philosophers could take it as certain that space has three dimensions, or that the laws of nature are not probabilistic—they could take these things as not needing empirical evidence—although were they more methodologically sophisticated they could accept that these certainties could conceivably be undermined.

Perception raises similar possibilities. Foundationalist epistemologists used to take the ultimate perceptual evidence as self-justifying, giving the bedrock of certainty on which more or less wobbly theories can rest. But this is a mad doctrine, something no one ever believed except when their philosophy maneuvered them into it. We can keep a hold on the primacy and fundamentality of perception with the saner assumption that what a person perceives normally gives them beliefs which they are not obliged to justify further unless certain weird circumstances obtain. One such circumstance would be if, whether or not they know it, the person finds herself in a situation in which her perceptual capacities are untrustworthy.

When, though? Which beliefs is it reasonable to take as certain, in the absence of what kinds of undermining facts? Writers who have made the defeasibility move have usually stuck to the examples. They argue simply that it is reasonable to trust your intuition that causes precede their effects or that the induction axiom is true of the integers. Or that it is reasonable to take the apparent readings of your instruments to be their actual readings if the perceptual situation seems normal. But there are many other ways of using the 'trust what seems clear to you, unless ...' strategy. It is often reasonable to take a hypothesis that seems implausible as not worth further investigation. It is often reasonable to react to an apparently insoluble difficulty with a theory as if it is indeed fatal to it. In these and other cases the reasonableness has to be defeasible: other things you believe or things you discover later may change the picture fundamentally. In many cases what is defeasibly reasonable is not an unqualified acceptance of a belief but a classification of a proposition as possible, worth considering, or unlikely. And it seems very plausible to me that the factors that make it reasonable to have such trust in oneself are thoroughly external, and the factors that would defeat the trust are quasi-external. That is, it makes sense to pay heed to intuitions that are in fact reliable indicators of the facts they claim to represent. And whether on a particular occasion the intuition ought to be heeded is a matter in part of what investigations the person has or has not carried out, what possibilities have and have not occurred to her, and so on. It is important that the story involves thoughts about what has *not* been considered, rejected, or inferred, in relation to what actually is the case, as much as the pattern of one's actual beliefs.

Defeasibility allows us to preserve a relative autonomy of perceptual belief and a priori intuition. But it is just a limiting case of departure from the principle of the absoluteness of evidence cited at the beginning of this section.

We can, more generally, insist on the autonomy of specific patterns of evidential relevance. That is, we can make an intuitively acceptable account of evidence, whose distinctions are not wiped out by the general fact of the potential relevance of anything to anything. In particular we can describe evidence in ways that make 'empirical', 'low-level', or 'non-theoretical' beliefs more relevant to its assessment than abstract or theoretical ones. (When considering whether a study of the incidence of some cancer in some population supported a theory of its causes one should take into account facts about the population and possible alternative causes of the data. But one should not consider theories of the origin of life or even the likelihood of alternative theories of the etiology of the cancer.) The motives for this might be methodological, as in the new experimentalism in the philosophy of science. Or ontological, as in Hacking's and van Fraassen's very different accounts of evidence. Whatever the reason, the fact is that with a little ingenuity we can defend the common-sense idea that in assessing evidence it is reasonable not to take account of many beliefs distant from the confirmational problem at hand. Reasonable, that is, as long as there are no positive grounds for operating with a wider and less manageable frame.[3]

5 Bayesianism to perceptual content

There is a natural reply here, from the internalist perspective. We are doing epistemology not psychology. The aim is to describe ways in which it is rational to form beliefs, whether or not they come easily to us. In fact, recent psychology teaches us how difficult it is for humans to follow even the simplest principles of inference. Probabilistic reasoning is a particular weak spot in the human head. And this is a particularly important example, because according to the most persuasive and best-worked-out internalist position, Bayesian confirmation theory, rational belief consists in updating one's probability assignments in accordance with new evidence. Recent work in Bayesian philosophy of science has made a good case that many intuitive methodological principles can be understood as approximations to sound probabilistic principles. (Moreover, the Bayesian point of view reveals the conditions under which intuitive principles are reasonable, in ways that cannot be extracted from naïve intuition.) From this point of view talk of intuition and defeasibility is beside the point, unless we can re-express it in

---

[3] For the defeasibility move with respect to a priori knowledge, see Boghossian ([1996]), Harman ([1996], Chs 1 and 4 of Bonjour ([1998]), Bealer ([1999]), and also Goldman ([1999a]). For the move with respect to perceptual evidence see McDowell ([1982]), Ch. 6 of Bonjour ([1985]), Plantinga ([1993]), and Audi ([1999]). For the general rehabilitation of intuition see the essays in DePaul and Ramsay ([1998]), especially Stich, Kornblith and Bealer. For theories of evidence that require a delicate adjustment with coherentism, see Glymour ([1980]), Hacking ([1983]), van Fraassen ([1989]), Woodward and Bogen ([1988]), Mayo ([1996]).

terms of approximations to methods of updating coherent sets of degrees of belief.

Valuable as the Bayesian contribution has been, recent anti-Bayesian work has also been very impressive and enlightening. Two strands in particular fit the themes of this article. First there is work on inductive learning. Given a stream of data that in its entirety conforms or does not conform to some general hypothesis, what data-to-generalization procedures will give the right Yes or No? The essential work here is by Kevin Kelly. Kelly's work shows that as we impose increasingly strict conditions on the kind of cognitive resources we bring to inductive reasoning (from unlimited power to Turing machine to finite state machine) an increasingly enlightening and plausible classification emerges, of the kinds of proposition that are amenable and resistant in various kinds of empirical confirmation. Formalizing the limits on our cognitive powers is as enlightening as formalizing our unattainable ideals. Moreover, well-performing inductive programs are not just cut down instances of Bayesian procedures. One very fundamental difference is the absence of prior probabilities (compensated for to some extent by a certain arbitrariness in the choice of program and background assumption). Kelly gives examples of data streams and true generalizations which cannot be confirmed by Bayesian updating from any priors, but which are easily verified by simple inductive procedures. (This complements examples by Earman and others of situations in which prior probabilities will not converge under conditionalization to true frequencies.) Kelly makes the provocative suggestion that Bayesianism fudges its confrontation with skepticism by assigning skeptical possibilities probability zero and hiding behind the assumption of countable additivity to force probability zero on things that clearly can occur.

The other relevant anti-Bayesian stream is provided by critics of conditionalization as a model of the response of a body of theory to new evidence. The problem is particularly acute when the evidence is perceptual. All coherentist epistemologies have a problem here. They have to explain why perceptual evidence has a particularly important role in changing our beliefs. Critics of epistemological holism have often produced examples of coherent sets of beliefs that maintain their coherence best if perception is ignored or dismissed. In the particular case of Bayesian epistemology the problem takes the form of finding a reason why the right response to a perceptual experience is some form of conditionalization. A traditional response is to take the experience to confer probability 1 on some proposition, prior to simple conditionalization. Or, more flexibly, to take the conferred probability to be less than 1 and to use Jeffrey's generalization of conditionalization. But the strategy is not, in either version, satisfactory all by itself. Most fundamentally, it does not tell us what is to count as perceptual evidence. If I bang my head and

with utter confidence take the resulting stars as angels, why is this proposition not to be given a probability commensurate with my conviction?

There are two responses to the problem: augment the theory or restrict its scope. The first has been explored by David Christensen. Christensen explores various ways of resolving the problem using resources internal to Bayesianism and finds them inadequate. He proposes that we understand the force of perceptual evidence on a two-factor model: an experience pushes a belief towards a particular probability with a certain force. (Perhaps the concussive stars I see suggest that if I give them credence I should give the existence of angels a high probability, but their rational force towards this target is very small. Or under different circumstances in which I can hardly believe my eyes they might push towards a much lower target probability with much greater force.) If anything like this model is right then a probabilistic epistemology will have to be augmented with principles which determine the epistemic status of perceptual experiences and their evidential relevance to propositions in a person's cognitive grasp. Perhaps, though, this is attempting to fuse inherently incomparable considerations. Timothy Williamson proposes that we restrict conditionalization to well-formed propositions whose epistemic status is well established, which can be given probability one and fed into simple conditionalization. (We conditionalize only on what is known.) Again the suggestion is that we take Bayesian principles as only part of epistemology, but now there is no complex interaction between the two parts. If we have a theory of perceptual knowledge we can use it to determine what propositions suggested by perception are known, and we can then safely conditionalize on them. If, as will very often be the case, we do not know precisely what it is that we perceptually know, all we can do is conditionalize when it seems reasonable, realizing that sometimes the result will be probabilistically correct but an epistemic mistake.

But why *is* perception a good source of evidence? Why is it reasonable to treat perceptual beliefs as special? Two characteristics of perception might seem relevant. The first is that when we allow perception to shape our beliefs we usually, though far from always, end up with beliefs that are coherent with one another and which fit into our attempts to get an explanatorily coherent system of beliefs. Indeed, internal to most coherent systems of beliefs are reasons why perception is generally reliable, describing causal chains from facts via the senses to belief. One worry with this answer derives from theory-ladenness. Given our tendency to shape our perceptions, we might fear that it amounts to no more than saying that if we believe that perception is reliable then it will rarely lead us to doubt that belief. Another relevant characteristic of perception is its phenomenal aspect. Very often when a person perceives they also have a perceptual experience: it looks or feels a certain way to them. Traditional epistemology makes much of the fact that these experiences,

inasmuch as we can separate them from the associated beliefs, are less sensitive to the overall pattern of our beliefs. This suits them to give unbiased testimony, to be evidence. The problem is that they are not themselves beliefs, cannot even be true or false, so it is not clear how they can play a part in inferences to beliefs. Much of the contemporary philosophy of perception can be seen as trying to deal with this problem. An influential approach, due largely to Christopher Peacocke, takes perceptual appearances to be equipped with their own set of almost-concepts. These are structural features of appearance that are linked to properties in the world, though they need not correspond to concepts that feature in the person's beliefs. Examples of such features are the angles between surfaces, or their patterns of occlusion. The presence of such features gives a perceptual appearance some of the structure and content that inference requires.

Any such account will owe an enormous debt, of explaining what forms inferences between non-conceptual quasi-propositional contents and real beliefs can take. But, even allowing the debt, it seems unlikely that non-conceptual content can be the whole story. For perceptual appearances also have features which do not correspond to environmental properties. For example the centeredness of the perceptual environment on the location of the perceiver is an intrinsic feature of perception that just happens to be illusory. Every child has to make the Copernican discovery that it is wrong. The existence of some systematic feature of perceptual appearances does not by itself show that any beliefs linked to this feature are reliable. Here some form of the first factor mentioned above seems necessary. We have to use our account of the world, incomplete and uncertain as it is, to tell us which aspects of perception can reasonably be taken as evidence.[4]

6 Finiteness to virtue

Elementary epistemology courses often begin with a sequence of inference methods of increasing riskiness: deduction, simple induction, inference to the best explanation. This sets up the valuable skeptic-undermining question 'how much risk of getting some false beliefs is it reasonable to take in order to have a chance of getting some valuable true beliefs?' (How do you weigh your

---

[4] For now-orthodox Bayesian epistemology, see Howson and Urbach ([1989]). An unorthodox Bayesianism is defended in Kaplan ([1996]). The tendency of conditionalization to drive probabilities towards the truth is discussed in Earman ([1992]). For human feebleness about probability see Stein ([1996]). On learning theory, see Kelly ([1996]). Works critical of Bayesian epistemology in various ways are Field ([1978]), Christensen ([1992]), Ch. 8 of Plantinga ([1993]), and Williamson ([1997]). For the need for a theory which applies to incoherent beliefs, see Foley ([1993]), especially Ch. 4. For the epistemic relevance of the contents of perception, see Lewis ([1980]), Ch. 6 of Dretske ([1981b]), Millar ([1991]) and Peacocke ([1992]). For the view that non-conceptual content cannot be relevant to perceptual evidence, see Davidson ([1989]).

aversion to error against your aversion to ignorance?) But there is a way in which it is deeply misleading. It conflates logical relations and cognitive processes. Given a problem in which what is wanted is a conclusion that is a logical consequence of given premises, what a person does is very rarely to churn out consequences of those premises. Usually they look ahead and think what strategies of deduction are promising. Given the more general problem of revising your beliefs with an aim to maximizing logical consistency and deductive closure, it is even less clear that any inferential process shadowing deduction plays much part. It certainly is not true that you should believe anything you discover to follow from your beliefs. Sometimes you'd be a fool not to abandon some of those prior beliefs instead. And even when it is clear what to add and what to abandon no deductive principle tells you this, any more than any such principle tells you what lines of deductive consequence are sensible or necessary to explore.

These observations are due to Gilbert Harman. They have most force when combined with a position defended by Christopher Cherniak. Cherniak argues that ideals such as deductive closure, which could be achieved by unlimited agents, are not only unachievable but undesirable as applied to real limited agents. An example of Harman's makes the point clearly. Suppose that a person who believes p suddenly encounters overwhelming evidence that not-p. Just for a moment she believes both p and not-p and then sensibly abandons p. In so doing she uses capacities that are essential for limited agents *and that are not available to unlimited agents.* For an unlimited agent would in the moment of inconsistency acquire all beliefs that followed from the contradiction, and thus be in an irrecoverable disaster. (That is not to say that there is no way out for an infinite agent, just that its ways of managing its beliefs would be almost inconceivably different from ours.)

From this point of view the line between deduction and induction becomes blurred. All people have beliefs that are unsatisfactory in many ways, and a rational person will revise her beliefs in the direction of logical consistency and explanatory coherence, inasmuch as the situation warrants. It is conceivable that some satisfactory Bayesian procedure for incorporating new information into a set of degrees of belief would provide a useful model here, though assumptions of logical coherence are so deeply built into Bayesianism that the model would need a lot of modification. It is also conceivable that some clear and useful analysis of the concepts of explanation and explanatory coherence would do the job, as Harman suspects. But we are a long way from seeing any solutions with the clarity that Harman and Cherniak describe the problems.

While it is not clear what the rational requirements on agents with messed-up beliefs (i.e. human beings) are, it is clear that there are ways and styles of thinking that are appropriate to the untangling of beliefs. Thinking ahead to see where a line of deduction might lead (vision: route-finding styles). Imagining

objections to a position (strategy: chess-playing styles). Waiting for a position to get thought out before embracing it (patience: fish-catching styles). Willingness to allow that a weak position might be true (courage: man-fighting styles). Realization that an apparently attractive position may have problems (caution: predator-avoiding styles). There is a name for these styles; they are epistemic virtues. There is now a flourishing subject of virtue epistemology. It aims first of all to establish that the process of belief formation can be evaluated in terms of a variety of characteristics of believers, then that considering these characteristics throws light on traditional questions about knowledge and rational belief. As far as I am concerned the existence and interest of epistemic virtues is well established, especially in the light of considerations about the rational response to limited cognitive powers and incoherent belief. The virtues that are most easily described do not seem to be specific to belief. As my parentheses above suggest, they apply to most intelligent activities. And, as far as I am concerned, the applications of virtue epistemology to problems posed in the traditional epistemological vocabulary have not been very impressive. The reason is that the strongest reasons for talking in terms of epistemic virtues are also reasons for not framing problems in that traditional vocabulary.

Consider, for example, definitions of justified belief in terms of virtues. Typically a virtuous epistemic process is defined as one that tends to produce true beliefs, and then a belief is justified if, roughly, it results from a virtuous process. 'Roughly' because the definition needs to be relativized to the circumstances under which the process tends to truth. But then it is hard to avoid three suspicions. First, that 'epistemic virtue' has just become another name for a reliable process, so that virtue epistemology collapses into a standard externalist position. Second, that the concept of justification is suppressing the prime characteristic of virtues, that they are plural and need balancing against one another. For example consider someone who makes a logically impeccable but completely wrong-headed deduction from an established belief. The result is true if the established belief is, but to take the trouble to believe it is to point her cognitive resources in an unprofitable direction. There is something wrong with her thinking. Is her belief justified? Probably yes on a traditional account.[5] Does it exhibit epistemic virtue? No, not if the concept is to have any distinctive bite at all. It exhibits the minor virtues of consistency and correct deduction and the major vice of strategic blindness. In many cases the pull between competing virtues will be subtler and harder to sum up. We will rarely want to sum up with a one-criterion judgement: justified

[5] But justification is at best dubious even on a traditional account if the case is slightly varied. Then the person has arrived at a belief which it would be best to abandon, together with the premises that led to it, but which she perversely though consistently adds to her stock of beliefs. It takes a wise eye to tell a reductio from a discovery.

or not justified. The third suspicion is that the attempt to mimic traditional epistemology sidelines our actual vocabulary of intellectual virtues. It becomes hard to see how to make use of the distinctive and competing characteristics of prudence, risk-taking, curiosity, common sense, and flair, if they must all add up to classifications of beliefs in terms of the much less rich language of justification and knowledge.

This is not to rubbish virtue epistemology. It is one of the interesting recent developments. But it is to argue that there are reasons why we have virtue terms in our everyday belief-evaluating vocabulary, and why we ought to search for sharpened versions for use in philosophy. These reasons derive from our need for belief structures with various desirable properties, and not just for piles of arbitrary truths. There are real opportunities here, which are likely to be missed if we waste them reconstructing old positions and re-solving old problems.[6]

7 The structure of belief: from holism to foreground/background

The epistemic status of a belief depends on its position in the entire body of that person's beliefs. Perhaps on its position in some body of other people's beliefs. This is so even on any foundationalist account which while justifying beliefs in terms of a one-way connection with experience gives a reasonable account of inductive evidence. For the force of any non-deductive link will depend not only on the evidence cited but also the fact that that there is no other contrary or complicating evidence available. So some weak holism will be common to almost any account of reasons for belief. There is a moral here. When a generation ago we escaped from a clutch of blinkering doctrines we did so in part by waving large contrary slogans: holism, fallibilism, anti-apriorism. But in fact the slogans are equally blinkering, as they assimilate different issues. One can be a holist and believe that beliefs need to be based on secure experiential evidence. One can be a fallibilist and leave room for the a priori. (The defeasibility move discussed above makes both of these defensible.) Moreover, there is a deep ambiguity in too-sweeping doctrines of global structure. Is the structure one that is found in the actual beliefs of reasonable people, or those of scientifically conscientious people, or those of fully rational people? If the latter is the rationality that of ideal epistemic agents or that of somewhat idealized agents to whom we humans can approximate? Given these

[6] For the different aims one can have in acquiring beliefs, see Ch. 1 of Goldman ([1986]) and Ch. 1 of Foley ([1993]0. For a minority view, see Stich ([1990]). For the situation of the finite reasoner, see Cherniak ([1989]) and Part I of Harman ([1999]). Rubinstein ([1998]) demonstrates that precise models of limited cognition are possible. On virtue epistemology, see Code ([1987]), Axtell ([1997]), Kornblith ([1983]), Sosa ([1991]), Greco ([1992]), Plantinga ([1993]), Zagzebski ([1996]), Hookway ([1999]).

questions, it is often not at all obvious what the claims and disputes are all about. In fact they are about many different things; I'll mention just two.

First, there are disputes about the relation between theory and evidence. Any coherentist account of evidence worth the name will insist that the evidence for a set of beliefs cannot be resolved into a set of independent evidential relations. For often the evidence is not sufficient for accepting A alone or B alone, but if both are accepted then each can fill the gaps between the evidence and the other. (Susan Haack uses the example of a set of crossword puzzle answers, no one of which would be reasonable unless it also fitted together with the others.) Most coherentist accounts will also insist that a hypothesis can be more certain than any single item of the evidence on which it is based. (Or, more carefully, more certain than any single item would be in the absence of the hypothesis.) Both of these claims are compatible with a variety of positions about the nature and force of perceptual evidence. They are compatible with an extreme view, which sees perceptual beliefs as simply beliefs among others, to be over-ruled if they do not cohere with theory. They are also compatible with opposite views, which make perceptual beliefs extremely solid and authoritative.

The coherentist attitude to evidence thus needs to be augmented with a view of the role of perception. There is a natural, almost inevitable, way of doing this. We should characterize perceptual evidence in terms of those actual connections between states of affairs and states of mind which have the properties required for the states to be evidence. Two such properties are that the states of mind reliably track the states of affairs and that they be relatively immune to influences from one another and other states of mind. These are both properties that can only be ascribed on the basis of information about the actual relations between states of mind and states of the world. There are many kinds of perceptual content that could turn out to be evidentially relevant, depending in part on whether they track facts and resist other states. Some sense modalities might not qualify as evidence, and some unconscious states might qualify. Filling in the blanks in a coherentist account of evidence thus leads to a particular kind of naturalism.

A similar conclusion emerges from the second dispute, the tunability of reasoning. A moderately strong coherentism will argue that the methods we use to support our beliefs are not fixed but are themselves part of the evolving body of doctrine. At the very least our ideas about what counts as a good explanation, and thus what counts as an explanatorily coherent body of beliefs, will itself change as science develops. A stronger position would urge that more fundamental factors are shaped by the pressure of the body of beliefs. For example, questions about the role explanatory force plays in inductive reasoning, or about the degree to which one ought to be conservative in changing ones beliefs, might be taken not as answerable a priori but as matters to be determined in the light of one's current pattern of belief.

Suppose belief formation is shaped by belief. How? Suppose the shaping is normative, a matter of how we should rather than how we do think. Then one might suppose that there were meta-principles that determined how method should change given belief. But this would amount to a collapse to a single a priori methodology.

The general connection between normative force and a priori principles can be blocked in several ways. There might be arbitrary or underdetermined or in principle indescribable aspects to the process that obstruct the meta-principles. A different kind of obstruction arises out of considerations of human finiteness. Any general meta-principles would surely be too complex to be followed by creatures of limited intelligence and memory. If stated we would not understand them, and if understood we would not follow them.

There is a very general point here. I shall call it the AEA point. A traditional understanding of reasons for belief finds a some-all-all pattern: there are principles that for all situations and all beliefs arising in those situations determine the rational changes of belief. A naively relativist understanding finds an all-all-some pattern: given any circumstances and any beliefs there are principles that regulate change of those beliefs in those circumstances. But the truth is neither of these but a more subtle all-some-all pattern: in any circumstances there are principles which govern all beliefs that arise in those circumstances. Circumstances here include general structural features of a person's beliefs and aspects of their material conditions that they may not be aware of. These are for the most part not represented in the person's explicit beliefs, and they determine principles of reasoning in ways that are evaluated, externalistically, in terms of their tendency to produce desirable principles. (Understanding how these tendencies operate may lead us to beware or compensate for some undesirable aspects.) But given circumstances and principles the evaluation of changes of belief is largely internalistic, in terms of standards that an agent can apply to his own knowledge of his own beliefs. Internalism/externalism, finite versus ideal, and defeasibility come together here. For finite agents it is inevitable that background circumstances determine externalistically principles which are normative with respect to a more manageable foreground.[7]

8 Conclusion: from social networks to the biology of science

There are many themes here, and many important connections between

---

[7] The current awareness that coherentism paints with much too broad a brush is due largely to Bonjour, Plantinga and Haack. See Bonjour ([1985]), Plantinga ([1993]), Haack ([1993]), especially Chs 4, 5, 6. Philosophers of science have been making similar points, see Glymour ([1980]), Mayo ([1996]). For a serious attempt at making coherentism work, see Harman ([1986]).

themes. Before picking out some particularly significant ones, let us step back and ask what use the scientific enterprise has for the vocabulary of epistemology. Science wants to be able to present some of its conclusions as knowledge; it wants to argue that many non-scientific or pseudo-scientific beliefs are unjustified; it needs to be able to evaluate the evidential support for conjectures and possible orthodoxies. For all these purposes there is an advantage to using terms that are continuous with those used when people evaluate everyday beliefs. After all there is little force to a criticism of a non-scientific belief in terms of criteria that seem to have been invented just for the purpose.

In the early days of science, which were by no accident also the early days of epistemology, the ambition was for a set of beliefs which could largely supplant ignorance and superstition, and which each person could master. So the epistemological vocabulary was to be available to all and to be employed from a neutral standpoint from which the advantages of science over ignorance could be adjudicated. Our situation and ambitions now are different. Each scientist is a lay-person with respect to most of science; no person can master all of scientific doctrine, and must be content with some common-sense approximation to much of it. It is very far from obvious that one epistemic vocabulary can give all that we need in this situation. The operating manual for the belief system must have several loosely linked chapters. One is directed at people accumulating everyday beliefs. Most likely the ideas needed for keeping one's own beliefs in order are much the same as those needed for evaluating those of others. Another is directed at people operating within a scientific context. Most likely the tension between what is ideally required and what can be expected of a mere human is at its greatest here. Another is directed at non-scientific assessment of expert opinions. Most likely the contrast between the point of view of the person who has the belief and the person who is assessing their grounds for it is very great here. And there are others.

Different strands in contemporary epistemology seem appropriate to different of these tasks. To that extent apparently competing theories may not always be real rivals. In particular, disputes involving the central concept of justification may very often lack substance, since the criteria for acceptability of a belief are so different for different purposes. The emphasis in much recent work on the third-person point of view makes sense in terms of the shift to an acceptance of the inevitability of dependence on one person's beliefs on another's expertise. Epistemic virtues also make sense in these terms. If all agents played the same epistemic roles then there would be little need for more than a single epistemic virtue of reasonableness. But when roles vary, and different people are called on to supply beliefs of a variety of kinds, which are inputs to those of others in a variety of ways, a variety of desirable and undesirable epistemic characteristics appears. The AEA point also finds its place. Crucial among the circumstances which tune the principles and criteria

relevant to a particular person's beliefs—and the virtues which it is important for her to exhibit—are facts about the use that others make of her reports, and the use that she makes of those of others.

The ubiquitous naturalistic tendency also fits with a many-sided conception of epistemology. The obvious route to naturalism starts from taking science seriously in all matters and from the absence of a neutral a priori standpoint for judging beliefs. As remarked above this creates a suspicion that science has cooked the books in its favor. There is another route. We can create a particular science for epistemic purposes, whose ideas link with common sense and also with biology, psychology and indirectly physics. Such a science would search for a characterization of human knowers, which would do justice to everyday intuitions while construing knowledge as a natural phenomenon. I know where my keys are, particle physicists know whether neutrinos have mass, salmon know the streams that lead to their native rivers. To understand the relations between everyday, scientific, and animal knowledge one would have to understand and accept this theory, whose grounds should therefore be available to informed common sense, but whose implications would characterize the variety of human knowledge as an exercise of natural capacities. (Considerations of human social nature and of the bounds to human cognition will inevitably play a large role.) It is far from obvious that such a science is available. But, I believe, it is the target of much current epistemology. It is worth aiming for.[8]

Acknowledgements

References

Audi, Robert [1999]: 'Self-evidence', in James E Tomberlin (ed.), *Philosophical Perspectives, 13: Epistemology*, Oxford: Blackwell, pp. 205–28.

Axtell, Guy [1997]: 'Recent Work on Virtue Epistemology', *American Philosophical Quarterly*, **34**, pp. 1–26.

Bealer, George [1998]: 'Intuition and the Autonomy of Philosophy', in Michael R DePaul and William Ramsey (eds), *Rethinking Intuition*, New Jersey: Rowman & Littlefield.

[8] For social roles in the distribution of belief, see Craig ([1990]), and Part 2 of Goldman ([1999b]). Feminist themes, explored in Code ([1987]) are also relevant here. Kornblith ([1999]) defends knowledge as a biological property, a view that is opposed in Williams ([1992]).

Bealer, George [1999]: 'The A Priori', in John Greco and Ernest Sosa *(eds)*, *The Blackwell Guide to Epistemology*, Oxford: Blackwell, pp. 243–70.

Boghossian, Paul [1996]: 'Analyticity Reconsidered', *Noûs*, **30**, pp. 370–91.

Bonjour, Laurence [1985]: *The Structure of Empirical Knowledge*, Cambridge, MA: Harvard University Press.

Bonjour, Laurence [1998]: *In Defence of Pure Reason*, Cambridge: Cambridge University Press.

Burge, Tyler [1979]: 'Individualism and the Mental', in P French, T Uehling and H Wettstein *(eds)*, *Midwest Studies in Philosophy* **4**, pp. 73–121.

Cherniak, Christopher [1989]: *Minimal Rationality*, Cambridge, MA: MIT Press.

Christensen, David [1992]: 'Confirmational Holism and Bayesian Epistemology', *Philosophy of Science*, 59, pp. 540–57.

Code, Lorraine [1987]: *Epistemic Responsibility*, Hanover, NH: University Press of New England.

Craig, Edward [1990]: *Knowledge and the State of Nature*, Oxford: Oxford University Press.

Davidson, Donald [1989]: 'Meaning Truth and Evidence', in R Gibson and R Barrett *(eds)*, *Perspectives on Quine*, Oxford: Blackwell.

DePaul, Michael R and Ramsey, William *(eds)* [1998]: *Rethinking Intuition*, New Jersey: Rowman & Littlefield.

DeRose, Keith [1995]: 'Solving the Skeptical Problem', *Phil Rev*, **104**, pp. 1–50.

Dretske, Fred [1981a]: 'The Pragmatic Dimension of Knowledge', *Philosophical Studies*, **40**, pp. 363–78.

Dretske, Fred [1981b]: *Knowledge and the Flow of Information*, Cambridge, MA: MIT Press.

Earman, John [1992]: *Bayes or Bust?*, Cambridge, MA: MIT Press.

Field, Hartry [1978]: 'A Note on Jeffrey Conditionalization', *Philosophy of Science*, **45**, pp. 361–67.

Foley, Richard [1993]: *Working Without a Net*, New York: Oxford University Press.

Foley, Richard [1998]: 'Rationality and Intellectual Self-trust', in Michael R DePaul and William Ramsey *(eds)*, *Rethinking Intuition*, New Jersey: Rowman & Littlefield.

Glymour, C [1980]:*Theory and Evidence*, Princeton: Princeton University Press.

Goldman, Alvin [1986]: *Epistemology and Cognition*, Cambridge, MA: Harvard University Press.

Goldman, Alvin [1999a]: 'Internalism Exposed', *Journal of Philosophy*, **97**, pp. 271–93.

Goldman, Alvin [1999b]: 'A Priori Warrant and Naturalistic Epistemology', in James E Tomberlin *(ed.)*, *Philosophical Perspectives, 13: Epistemology*, Oxford: Blackwell, pp. 1–28.

Goldman, Alvin [1999c]: *Knowledge in a Social World*. Oxford: Oxford University Press.

Greco, John [1992]: 'Virtue Epistemology', in Jonathan Dancy and Ernest Sosa *(eds)*, *A Companion to Epistemology*, Oxford: Blackwell.

Haack, Susan [1993]: *Evidence and Inquiry*, Oxford: Blackwell.

Hacking, Ian [1983]: *Representing and Intervening*, Cambridge: Cambridge University Press.

Harman, Gilbert [1986]: *Change in View*, Cambridge, MA: MIT Press.

Harman, Gilbert [1996]: 'Analyticity Regained', *Noûs*, **30**, pp. 392–400.

Harman, Gilbert [1999]: *Reasoning, Meaning, and Mind*, Oxford: Oxford University Press.

Hookway, Christopher [1999]: 'Epistemic Norms and Theoretical Deliberation', *Ratio*, 12, pp. 380–98.

Howson, Colin and Urbach, Peter [1989]: *Scientific Reasoning: The Bayesian Approach*, La Salle, IL: Open Court.

Kaplan, Mark [1996]: *Decision Theory as Philosophy*, Cambridge: Cambridge University Press.

Kelly, Kevin T [1996]: *The Logic of Reliable Inquiry*, New York: Oxford University Press.

Kirkham, Richard [1984] 'Does the Gettier problem rest on a mistake?' *Mind*, **93**, pp. 501–13.

Kornblith, Hilary [1983]: 'Justified Belief and Epistemically Responsible Action', *Philosophical Review*, **92**, pp. 33–48.

Kornblith, Hilary [1999]: 'Knowledge in Humans and Other Animals', in James E Tomberlin *(ed.)*, *Philosophical Perspectives, 13: Epistemology*. Oxford: Blackwell.

Lewis, David [1980]: 'Veridical Hallucinations and Prosthetic Vision', *Australasian Journal of Philosophy*, **58**, pp. 239–49.

Lewis, David [1996]: 'Elusive Knowledge', *Australasian Journal of Philosophy*, **74**, pp. 549–67.

McDowell, John [1982]: 'Criteria, Defeasibility, and Knowledge', *Proceedings of the British Academy*, **68**, reprinted in Jonathan Dancy *(ed.)*, *Perceptual Knowledge*, Oxford: Oxford University Press [1988].

Mayo, Deborah [1996]: *Error and the Growth of Experimental Knowledge*, Chicago: Chicago University Press.

Millar, Alan [1991]: *Reasons and Experience*, Oxford: Oxford University Press.

Peacocke, Christopher [1992]: 'Scenarios, Concepts, and Perception', in Tim Crane *(ed.)*, *The Contents of Experience*, Cambridge: Cambridge University Press.

Plantinga, Alvin [1993]: *Warrant and Proper Function*, Oxford: Oxford University Press.

Rubinstein, Ariel [1998]: *Modelling Bounded Rationality*, Cambridge MA: MIT Press.

Shope, R K [1983]: *The Analysis of Knowing*, Princeton: Princeton University Press.

Sosa, Ernest [1991]: *Knowledge in Perspective*, Cambridge: Cambridge University Press.

Stein, Edward [1996]: *Without Good Reason*, New York: Oxford University Press.

Stein, Gail [1976]: 'Skepticism, Relevant Alternatives, and Deductive Closure', *Philosophical Studies*, **29**, pp. 249–61.

Stich, Stephen [1990]: *The Fragmentation of Reason*, Cambridge MA: MIT Press.

Stich, Stephen [1998]: 'Reflective Equilibrium, Analytic Epistemology, and the
   Problem of Cognitive Diversity', in Michael R DePaul and William Ramsey
   (*eds*), *Rethinking Intuition*, New Jersey: Rowman & Littlefield.

Unger, Peter [1975]: *Ignorance: A Case for Scepticism*, Oxford: Oxford University
   Press.

van Fraassen, Bas [1989]: *Laws and Symmetry*, New York: Oxford University Press.

Vogel, Jonathan [1999]: 'The New Relevant Alternatives Theory', in James E
   Tomberlin (*ed.*), *Philosophical Perspectives, 13: Epistemology*, Oxford: Blackwell,
   pp. 155–80.

Welbourne, Michael [1986]: *The Community of Knowledge*, Aberdeen: Aberdeen
   University Press.

Williams, Michael [1977]: *Groundless Belief*, Oxford: Oxford University Press.

Williams, Michael [1992]: *Unnatural Doubts*, Oxford: Blackwell.

Williamson, Timothy [1997]: 'Knowledge as Evidence', *Mind*, **106**, pp. 717–42.

Woodward, James and Bogen, James [1988]: 'Saving the Phenomena', *Philosophical
   Review*, **99**, no. 3, pp. 303–52.

Zagzebski, Linda [1996]: *Virtues of the Mind*, Cambridge: Cambridge University Press.

3

The Present State of the Scientific Realism Debate

Stathis Psillos

The unique attraction of realism is the nice balance of feasibility and
dignity that it offers to our quest of knowledge . . . . We want the mountain
to be climbable, but we also want it to be a real mountain, not some sort of
reification of aspects of ourselves (Wright [1988], p. 25).

ABSTRACT

In this survey article I try to appraise the present state of the scientific realism debate
with an eye to important but hitherto unexplored suggestions and open issues that need
further work. In Section **2**, I mostly focus on the relationship between scientific realism
and truth. In Section **3**, I discuss the grounds for realists' epistemic optimism.

1 Introduction

There was at one time a feeling in the philosophy of science community that
the scientific realism debate had run out of steam. Fine went so far as to declare
that 'realism is well and truly dead' ([1986a], p. 112) and to write the obituary
of the debate, *aka* the Natural Ontological Attitude. Fortunately, this line of
argument has failed to persuade many philosophers, who still think that the
scientific realism debate has a glorious past and a very promising future. In the
last dozen years alone there have been a number of books which cast a fresh
eye over the issue of scientific realism, such as those by Suppe ([1989]),
Putnam ([1990]), Almeder ([1992]), Wright ([1992]), Kitcher ([1993a]),
Aronson, Harré and Way ([1994]), Brown ([1994]), Laudan ([1996]), Leplin
([1997]), Kukla ([1998]), Trout ([1998]), Cartwright ([1999]), Giere ([1999]),