Philosophical 'Intuitions' and Scepticism about Judgement

Timothy WILLIAMSON[†]

ABSTRACT

1. What are called 'intuitions' in philosophy are just applications of our ordinary capacities for judgement. We think of them as intuitions when a special kind of scepticism about those capacities is salient. 2. Like scepticism about perception, scepticism about judgement pressures us into conceiving our evidence as facts about our internal psychological states: here, facts about our conscious inclinations to make judgements about some topic rather than facts about the topic itself. But the pressure should be resisted, for it rests on bad epistemology: specifically, on an impossible ideal of unproblematically identifiable evidence. 3. Our resistance to scepticism about judgement is not simply epistemic conservativism, for we resist it on behalf of others as well as ourselves. A reason is needed for thinking that beliefs tend to be true. 4. Evolutionary explanations of the tendency assume what they should explain. Explanations that appeal to constraints on the determination of reference are more promising. Davidson's truth-maximizing principle of charity is examined but rejected. 5. An alternative principle is defended on which the nature of reference is to maximize knowledge rather than truth. It is related to an externalist conception of mind on which knowing is the central mental state. 6. The knowledge-maximizing principle of charity explains why scenarios for scepticism about judgement do not warrant such scepticism, although it does not explain how we know in any particular case. We should face the fact that evidence is always liable to be contested in philosophy, and stop using talk of intuition to disguise this unpleasant truth from ourselves.

1. Scepticism about judgement.

When contemporary analytic philosophers run out of arguments, they appeal to intuition. Intuitiveness is supposed to be a virtue, counterintuitiveness a vice. It can seem, and is sometimes said, that any philosophical dispute, when pushed back far enough, turns into a conflict of intuitions about ultimate premises: 'In the end, all we have to go on is our intuitions'. Yet analytic philosophy has no agreed or even popular account of how intuition might work, no accepted explanation of the hoped-for correlation between our having an intuition that P and its being true that P. Since analytic philosophy prides itself

† Faculty of Philosophy, University of Oxford. Email: timothy.williamson@philosphy.oxford.ac.uk

on its rigour, this blank space in its foundations looks like a methodological scandal. What is intuition? Why should it have any authority over the philosophical domain?

On closer inspection, putative examples of intuition tend to dissolve. Its powers are easily exaggerated under the pressure of debate. I once heard a professional philosopher argue that persons are not their brains by saying that he had an intuition that he weighed more than three pounds. Surely there are better ways of weighing oneself than by intuition. But such inapposite appeals to intuition are not idiosyncratic misjudgements. They are signs of a deeper elusiveness.

In contemporary discussion of the philosophical use of intuition, the canonical example is the judgement that the subject in a Gettier case does not know. Edmund Gettier (1963) constructed cases in which someone has a justified false belief that P, has competently deduced the true conclusion that Q from the premise that P, and on that basis has a justified belief that Q (in some relevant sense of 'justified'). Nevertheless, most philosophers judge, the subject, despite having a justified true belief that Q, does not know that Q. Since Gettier's article first appeared, the general consensus has been that such cases refute the traditional analysis of knowledge as justified true belief. But doubts have been raised about the soundness of the underlying methodology. For example, there is some evidence that the judgement that the subject does not know is not uniform across cultures (Weinberg, Stich and Nichols 2001). Why should we attach so much weight to a mere intuition? Why should our having it be good evidence for its truth?

The term 'intuition' tends to suggest something brute and simple. Yet, in judging that the subject in a Gettier case does not know, we employ a complex of cognitive capacities. Some of them are required because we are performing a kind of thought experiment. Gettier's cases are imaginary. To use one against the traditional analysis of knowledge, we must make modal judgements of at least two sorts. First, we must make a judgement of possibility: the case could have occurred (as described neutrally, without use of 'know' or cognate terms). Second, we must make a counterfactual conditional judgement: if the case had occurred, then the subject would have had a justified true belief that Q without knowing that Q. Together, these two judgements entail another judgement of possibility: that there could have been someone who had a justified true belief that Q without knowing that Q. That result is inconsistent with the principle that, necessarily, one knows that Q if and only if one has a justified true belief that Q, and thereby forces the rejection of the analysis of knowledge as justified true belief, because analyses are understood as implying statements

of necessary and sufficient conditions. Thus one way of objecting to the use of such cases to establish epistemological conclusions is to raise doubts about the kinds of modal thinking on which it rests. How can we tell whether something non-actual is possible, or what would have been the case if it had occurred? Is what is peculiarly problematic about intuition its use in thinking about counterfactual possibilities?

The epistemology of modal thought is notoriously problematic. Neither conceivability nor consistency is a safe guide to possibility. Nevertheless, there are reasons to think that these problems miss the heart of concerns about intuition. For a start, typical Gettier cases are quite mundane physical and psychological possibilities. They hardly stretch our capacity for modal thinking to the limit or beyond in the way that some bizarre, far-out thought experiments seem to do in metaphysics. Although the traditional analysis of knowledge may itself be intended to apply to all metaphysical possibilities whatsoever, both mundane and far-out, a mundane counterexample is sufficient, though not necessary, to refute it. For many practical purposes, it is advantageous to be capable of recognizing some truths about mundane possibilities. I might recognize that the path I am on is dangerous by recognizing that I could easily have failed to duck, and that if I had failed to duck then the falling stone would have hit me. In general, when we have at least a partial ability to verify or falsify the thoughts that R and that S, we typically also have in consequence at least a partial ability to verify or falsify the thoughts that it could easily have been that R and that if it had been that R then it would have been that S. We can project a given cognitive capacity for dealing with the actual at least some way into the possible, even though it is much clearer *that* we can do so than *how* we can do so. Our capacity for somewhat reliable modal and counterfactual thought is hardly surprising, for we cannot know in advance exactly which possibilities are or will be actual. We need to make contingency plans. Moreover, thought about counterfactual cases often casts light on actual cases: even if causation cannot be defined in counterfactual terms, we can learn much about actual causal connections by thinking through counterfactual cases. In practice, the only way for us to be cognitively equipped to deal with the actual is by being cognitively equipped to deal with a variety of contingencies, many of them counterfactual. Of course, a philosopher may raise specific sceptical doubts about the reliability of modal thinking, but that concern is not the same as the general concern about the reliability of intuition.

We can test the last point by supposing that a Gettier case is actual, and that we know about it in adequate detail by standard empirical means. Indeed, we can probably discover actual examples by trawling through history; failing

that, we can bring about an actual Gettier case by playing a hoax on someone, if we can be bothered. In reacting to the actual case, we can dispense with the modal thinking about hypothetical Gettier cases. When we classify the case in epistemic terms, we need not first judge that it has some further features, then make the modal judgement that, necessarily, any case with those further features is a case of justified true belief without knowledge, and finally infer the conclusion that this is a case of justified true belief without knowledge. Rather, in the light of our first-hand experience of the case, we can make that episte-mological judgement without taking any detour through modal judgements about hypothetical Gettier cases. Would this more direct procedure satisfy those who doubt the Gettier intuition? Hardly. They will probably reply that intuition is still required to classify the actual case as one of justified true be-lief without knowledge. Yet the non-modal procedure involves the same ca-pacity to classify empirically encountered cases with respect to knowledge as we use when, for example, we classify a politician as not knowing the truth of his claims about terrorists.[1]

The idea that intuition is required even to classify actual cases is elicited by many appeals to such cases in metaphysics. For example, some revisionary metaphysicians deny that there are mountains.[2] They deny a proposition of the sort that G. E. Moore defended in his defence of common sense (Moore 1925). They concede that microscopic particles exhibit collective behaviour in the presence of which it is usual to believe that a mountain is present, but they classify that belief as false. They hold that although the ordinary use of the word 'mountain' has some utility, because it registers genuine discriminations between different sorts of situation in which different sorts of action are ap-propriate, it also embodies a mistaken metaphysical theory as to what the dif-ference between those sorts of situation consists in (of course, sceptics who

---

[1] Objection: Modal and counterfactual thinking is applied here to what *would* happen *if* philosophers encountered real life Gettier cases. Reply: Scepticism about this application is not crucial to scepticism about the Gettier intuitions. This reinforces the point that such scepticism focusses on the application of epistemic rather than modal and counterfactual concepts. Another objection: Nozick (1981, 172-96) analyses knowledge in counterfactual terms; on his view, any judgement about knowledge implicitly involves judgements concerning counterfac-tual conditionals. Reply: His analysis does not make philosophical judgements about knowl-edge any more modal than non-philosophical ones are. Anyway, sceptics about epistemological intuition make no appeal to counterfactual analyses of knowledge. After all, the way in which Nozick reaches his conclusions exemplifies the very methodology about which they are scep-tical. Nor would they regard their scepticism as undermined by growing evidence that counter-factual analyses of knowledge are incorrect (Williamson 2000, 147-63). Their scepticism is intended to get its grip irrespective of whether knowledge is a modal matter.
[2] Van Inwagen 1995 and Horgan 1996 defend similar views. The text presents a typical view without attempting to follow any one metaphysician in detail.

doubt that there are really any mountains may also be committed to doubting that there are really any words or beliefs, but let us ignore such complications for the time being, just as the sceptics tend to do). The claim that there are no mountains is usually regarded as counterintuitive. Even its proponents may concede that it is counterintuitive, but argue that the cost to intuition is worth paying for the overall gain in simplicity, strength, logical coherence and consonance with the results of the natural sciences that they attribute to their total metaphysical system, of which the claim is a consequence. If their system also entails that there could not have been mountains, then it contradicts the modal intuition that there could have been mountains. But far more striking is the contradiction between the non-modal claim that there are no mountains and the common sense judgement that there are mountains, for example in Switzerland. Whether or not they themselves agree that there are no mountains, many contemporary metaphysicians would think it philosophically naive to dismiss the revisionary metaphysical system out of hand by appealing to our elementary geographical knowledge that there are mountains in Switzerland. They may concede that perception, memory and testimony contribute much to the judgement, but insist that intuition also plays a part in judging that what there is in Switzerland amounts to mountains, and that we cannot assume that such intuitions are reliable. Thus doubts about intuition also arise for non-modal judgements. Even the straightforward perceptual judgement 'Those are mountains' (pointing in the Alps) is subject to such doubts, so they extend to what would ordinarily be classified as judgements of perception rather than of intuition.

Examples involving empirically encountered cases suggest that scepticism about intuition consists not in scepticism about a special kind of judgement but in a special kind of scepticism about any judgement. That scepticism does not target the most distinctive features of perception, memory, testimony, or inductive or deductive inference. Rather, it targets our practices of applying concepts in judgement. Let us call such scepticism *scepticism about judgement*. It does not question the existence of an external world to which we are causally related in the ways appropriate to perception, for example (at least, not until the concepts of causation and perception themselves come under scrutiny). Indeed, many sceptics about judgement are naturalists; their rhetoric is frequently scientistic. They present themselves as identifying ways in which our conceptual practices need, or may need, revision in the light of scientific advances that those practices, unsurprisingly, failed to anticipate. They doubt that we should go on in the same way.

A sceptic about judgement need not advocate scepticism about all judgements; total scepticism about judgement, if thoroughly applied, would result

in total intellectual paralysis. The term 'sceptic about judgement' will be applied to those who are sceptical in the way just described about some contextually relevant judgements. Whether their sceptical arguments generalize further than they would like is another matter.

Sceptics about judgement question in particular our standards for applying ordinary concepts in experience, for example the concept of a mountain, or the concept of knowledge. Of course, ordinary modal concepts are not immune to such attacks. Sceptics about judgement may criticize our standards for applying the concept of possibility, or of the counterfactual conditional. But such criticisms are just more cases of scepticism about judgement; they are neither the only cases nor uniquely central ones. There is a tendency to call judgements 'intuitive' in a given context, whether or not they are modal in content, when the form of scepticism that arises most saliently for them in that context is scepticism about judgement. In that sense, even a perceptual judgement may count as intuitive. Similarly, the existential judgement 'There are mountains' may be considered intuitive even though it is inferential, derived from the perceptual demonstrative judgement 'Those are mountains' by a step of existential generalization. In what follows, the word 'intuition' will be used in that loose way, without any purported reference to a mysterious faculty of intuition.

Like other sceptics, sceptics about judgement construct scenarios to explain how we could come to make the judgements at issue even if they were false. The debunking explanation is supposed to convince us that massive deception is at least possible. The scenarios for scepticism about judgement are often distinctive in attempting to verify the scientific image of the world while falsifying the manifest image, the common sense view of the world or what passes for such in our culture. Sometimes they allow that the ability to use the key terms of ordinary language (such as 'mountain') in the ordinary way confers some evolutionary advantage, because it helps us communicate to each other genuine but perhaps misdescribed differences between different situations. The disposition to apply such terms immediately on the basis of casual observation contributes to practical efficiency. Such unreflective discriminations have survival value in harsh environments, where quick decisions are needed. If our ancestors could not have made them before discovering the true theory of reality, we should not be here. Although the physical theory embedded in our intuitions has to be approximately correct in its predictions over a limited range of practically important cases, we do not expect it to match or even resemble the true physics in its representation of the underlying reality. Why should we expect intuition to do much better elsewhere? The cheapest, fastest and easiest conceptual route for us to making some useful discrimina-

tions may run through dirty ways of thought that presuppose a false but convenient metaphysics.

In other cases, sceptics may regard a conceptual practice as of merely local value, or even as doing more harm than good. For example, if standards for applying the term 'know' vary radically with cultural background, then an evolutionary explanation of my current standard is less plausible. The sceptic might tell a different and more sociological story about the cultural role of knowledge ascriptions that detaches them from their truth-conditions. The story might imply that such ascriptions nevertheless fulfil a positive social function to which their cultural variability adapts them. But we can also envisage more sinister stories, on which they somehow serve as instruments of intellectual repression.

Traditional sceptics argue that we do not know that we are not in a sceptical scenario; they do not argue that we actually are in such a scenario, for their point is that we cannot know what our situation really is. For them, the claim that we are in the common sense scenario is no better in epistemic status, but also no worse, than the claim that we are in the sceptical scenario. By contrast, many sceptics about judgement argue that we actually are in their sceptical scenario, for example in which there are no mountains. If they hold that we can recognise that their argument is sound, they must also hold that we can deduce that we are actually in their sceptical scenario. That involves them in no immediate inconsistency, for their scepticism is intended to be partial; they might compare it to scepticism about superstition. They present their views as superior to intuitive judgements in compatibility with the results of the natural sciences. They take for granted that those results have some positive epistemic status. Indeed, they often treat them as scientific knowledge. They feel a crisis of confidence in common sense, but not in scientific method.

Sceptics about judgement need not puritanically insist that nobody should ever say things like 'There are mountains in Switzerland'. At least some of their debunking explanations imply that in many everyday contexts those are good, useful things to say: outside the metaphysics seminar, utterances of 'There are mountains in Switzerland' tend to have more desirable effects than do utterances of 'There are no mountains in Switzerland'. Discovering the true theory of metaphysics will not change that. Even revisionary metaphysicians can continue to say such things, just as they can continue to say things like 'The sun will rise at 6 a.m. tomorrow'. But, they hold, those things are not strictly and literally true: the sun will not strictly and literally *rise* at 6 a.m. tomorrow, and there are not strictly and literally any *mountains* in Switzerland. If we want to think what is really true, we must think with the learned; but for

many purposes it is enough to say what is *to all appearances* true, and speak with the vulgar. We can live most of our lives on the basis of a fiction, revisionists may say; only when we take a more scientific attitude are we forced to recognize the fiction for what it is.

Is intuition anything more than the last resort of dogmatic conservativism, in its desperate attempt to hold back the forward march of scientific and metaphysical progress? We may wonder how sceptics about judgement can prevent their scepticism from spreading as far as the sciences themselves. For it infects standard perceptual judgements, on which the natural sciences systematically depend: microscopes, telescopes and other scientific instruments enhance ordinary perception but do not replace it, for we need ordinary perception to use the instruments. Moreover, when scientists judge that a given complex body of evidence of various kinds supports one theory against another, what is the status of their judgement? In applying concepts of epistemic appraisal they are not immune to scepticism about judgement. For instance, they are vulnerable to it when they judge that empirical evidence tells against the reliability of intuition. In practice, sceptics about judgement are often sceptical about only a few intuitions or concepts at a time, but the underlying forms of argument are far more general. We may suspect that scepticism about judgement is a bomb that, if it detonates properly, will blow up the bombers and those whom they hope to promote together with everyone else. But it does not follow that we can dismiss scepticism about judgement as self-defeating. That the revolutionary movement would be incapable of establishing a stable new government of its own does not show that it is incapable of bringing the old government down. At worst, sceptics about judgement are troublemakers who put on the table arguments that we find powerful and in need of a proper response, irrespective of their dubious motives for putting them there.[3]

2. Intuitions and appearances.

Different forms of scepticism distinguish themselves from each other by questioning some things while leaving others unquestioned. The sceptic about induction grants that all emeralds observed so far were green, in order to question the distinctively inductive step to the conclusion that all emeralds will always be green. The sceptic about deduction grants the premises that if P then Q and that P of an inference by modus ponens, in order to question the distinctively deductive step to its conclusion that Q. The sceptic about testimony grants that someone has said that it was raining, but questions whether she

[3] Compare Feyerabend 1978, 143.

spoke the truth. The sceptic about memory grants that my experience is as of remembering that it was raining, but questions whether I really remember that it was raining. Scepticism about perception grants that my experience is as of seeing that it is raining, so that it visually appears to me that it is raining, but questions whether the experience is veridical. In each case, the sceptic concedes an evidential base, but accuses us of going illegitimately beyond it. For the sceptic about judgement, often the only evidential base to hand short of the disputed proposition itself is the conscious inclination to assent to that proposition, to make the judgement.

If scepticism about judgement is treated by analogy with scepticism about perception, then its evidential base will be described as intellectual appearances, somehow analogous to perceptual appearances. George Bealer has defended just such an account of intuitions as intellectual seemings (Bealer 1998, 207; Bealer 2002, 73). Of course, they lack the rich phenomenology of perceptual seemings. In its perceptually appearing to one that something is the case, almost always much else perceptually appears to one too: that various things have various specific shapes and sizes, colours, sounds, tastes, textures, smells .... By contrast, in the moment of its intellectually appearing to one that something is the case, often little else intellectually appears to one. Perhaps mathematical intuition can have a rich phenomenology, even a quasi-perceptual one, for example in geometry: but the conscious inclination to judge that the subject in a Gettier case lacks knowledge is not like that. Any accompanying imagery is irrelevant. Nevertheless, the phenomenological disanalogy between perceptual seemings and intellectual seemings, in the sense of conscious inclinations to judgement, does not show that no epistemological analogy can be maintained.

Scepticism about perception typically narrows one's evidential base to one's present internal mental state. When I can see and hear and feel that it is raining, I might suppose my total evidence to include the fact that it is raining, available for assessing hypotheses, for example the hypothesis that the grass will grow.[4] But the sceptic about perception insists that I have available as evidence only the fact that it perceptually appears to me that it is raining, for sometimes what perceptually appears to me is not the case. From that fact about my present internal mental state I am challenged to reason legitimately outwards to the conclusion about my external environment that it really is raining. The sceptic about perception asks by what right I treat the fact that it perceptually appears to me that it is raining as good evidence that it is raining.

---

[4] The word 'fact' is used throughout for true propositions.

Scepticism about judgement narrows and internalizes our evidential base in a similar way without going as far as scepticism about perception, since it does not usually exclude mental states of other people or at other times. After reading Gettier's article, we might suppose our total evidence to include the fact that the subject in a Gettier case lacks knowledge, as evidence available for assessing hypotheses, such as the hypothesis that justified true belief is knowledge. But the sceptic about judgement insists that we have available as evidence only the fact that it intellectually appears to us (perhaps the members of a restricted group) that the subject in a Gettier case lacks knowledge, in other words, the fact that we are consciously inclined to judge that the subject lacks knowledge, for sometimes what intellectually appears to us is not the case. From that fact about our internal mental states we are challenged to reason legitimately outwards to the conclusion that the subject in a Gettier case really does lack knowledge. The sceptic about judgement asks by what right we treat the fact that it intellectually appears to us that the subject in a Gettier case lacks knowledge, the fact that we are consciously inclined to assent to that proposition, as good evidence that the proposition is true.

How does sceptical pressure psychologize our conception of our evidence? Given that it is rational to proportion one's degree of credence in a proposition to its probability on one's evidence, those who assume that one must always be in a position to know what rationality requires of one will also tend to suppose that one must always be in a position to know what one's evidence is. In that sense, they require an *operational* standard of evidence. They will therefore be reluctant to admit any differences in evidence between sceptical scenarios and the corresponding non-sceptical scenarios, for typically, if one is in a sceptical scenario, one is in no position to know that one is not in the corresponding non-sceptical scenario, and is therefore in no position to know that one's evidence differs from what it would have been in that non-sceptical scenario. Since the fact that it is raining is not available as evidence in the sceptical scenario in which it merely perceptually appears to one that it is raining, the fact that it is raining is held to be equally unavailable as evidence in the non-sceptical scenario in which it really is raining, all one's perceptual systems are functioning normally and one seems to perceive that it is raining. Similarly, confronted with a philosophical example, it may appear as obvious to us that P as that the subject in a Gettier case lacks knowledge, even though it subsequently turns out to be false that P. For example, 'P' might stand for the proposition in a subtle semantic paradox that if John says 'Everything Mary says is true' then John speaks truly if and only if everything Mary says is true. Although we may later be in a position to know that it is untrue that P, and

therefore not obvious that P, nevertheless when we originally judged that P we were not in a position to know that it was untrue that P, or even that it was less obvious that P than that the subject in a Gettier case lacks knowledge, for we were not then in a position to know the results of extended philosophical discussion of the example. A standard of evidence whose application can involve contentious philosophizing is not operational in the intended sense. On any standard of evidence on which our evidence includes the fact that the subject in a Gettier case lacks knowledge, we are not always in the relevant sense in a position to know what our evidence is. Contrapositively, we are always in a position to know what our evidence is only if our evidence does not include the fact that the subject in a Gettier case lacks knowledge. Thus, for both perception and philosophical judgement, the demand for an operational standard of evidence drives even many non-sceptics to adopt the psychologized standard of evidence to which the sceptics appeal. By that standard, one's evidence is only the fact that it perceptually or intellectually appears to one that P, not the fact that P itself.

The result is the uneasy conception which many contemporary analytic philosophers have of their own methodology. They think that, in philosophy, our ultimate evidence consists only of intuitions. Under pressure, they take that not to mean that our ultimate evidence consists of the mainly non-psychological putative truths that are the contents of those intuitions. Rather, they take it to mean that our ultimate evidence consists of the psychological truths that we have intuitions with those contents, whether true or false.[5] That is, our ultimate evidence in philosophy amounts only to psychological facts about ourselves. Nevertheless, they do not want the psychological fact that we have an intuition that P to be perfectly neutral with respect to the non-psychological question whether P, for that would lead to scepticism about philosophy, with the possible exception of philosophy of mind and language. If we merely seek the best explanation of our having the intuitions, without any presumption in favour of their truth, we may find a psychological theory to explain them, but how are we to answer the questions about a mainly non-psychological universe that grip many metaphysicians and other philosophers? Perhaps intuitions about thought and language have a special epistemic status, because they help to constitute their own subject matter; but to generalize that claim to all intuitions in philosophy is to fall into a silly idealism. The nature of identity over

---

[5] A recent example is Brian Weatherson (2003, 27), who assumes that the argument from Gettier cases against the justified true belief analysis of knowledge has the premise 'Intuition says that Gettier cases are not cases of knowledge' rather than the simpler 'Gettier cases are not cases of knowledge'.

time, for example, is not a matter of thought or language: the question is how things persist, not how we think or say that they persist. In explaining why we have intuitions, analytic philosophy has a preference for explanations on which those intuitions come out true over explanations on which they come out untrue, but the justification for that preference remains unclear. Even if we have an intuition that the former sort of explanation is better than the latter, why should we give that intuition a special privilege over others by adopting a methodology that assumes its truth? That our ultimate evidence in philosophy consists of facts about intuitions and that explanations of those facts on which the intuitions come out true are better (*ceteris paribus*) than explanations on which they do not are themselves epistemological rather than psychological claims. When taken far enough, the psychologization of philosophical method becomes self-defeating. Psychologism is no more a psychological theory than the Pythagorean doctrine that everything consists of numbers is a mathematical theory.[6]

The search for a purely operational standard of evidence is in any case vain. It can be argued on quite general grounds that *no* standard of evidence is such that we are always in a position to know what our evidence is (Williamson 2000, 93-113; 147-83). The general argument will not be rehearsed here, but it is easy enough to gain some sense of the difficulty. Whatever Descartes thought, facts about one's own present consciousness are not always cognitively accessible to one. For example, how am I to judge whether my present conscious inclination to judge that P is strong enough to count as an intuition that P? If all such inclinations count, no matter how weak, then the category of intuition will lump together very weak and very strong intuitions: yet an adequately fine-grained theory of evidence must surely discriminate between very weak and very strong intuitions in evidential impact. If the strength of intuitions is taken into account, the evidence will be recorded in something like the form 'I have an intuition of strength s that P'. The strength parameter s will have to be specified according to some common scale, in order to permit the comparisons between the strengths of sometimes conflicting intuitions which the theory of evidence will need to make. But then there will be plenty of scope for misjudging the strength of one's intuitions. After all, philosophers defending a given position against opponents have a powerful vested interest in persuading themselves that the intuitions that directly or indirectly favour it are

---

[6] Pust 2001 argues carefully that the following principle is self-defeating: 'Aside from propositions describing the occurrence of her judgements, S is justified in believing only those propositions which are part of the best explanation of S's making the judgements that she makes'. This represents a change in position from Goldman and Pust 1998.

stronger than they actually are. The stronger those intuitions, the more those who appeal to them gain, both psychologically and professionally. Given what is known of human psychology, it would be astonishing if such vested interests did not manifest themselves in at least some degree of wishful thinking, some tendency to overestimate the strength of intuitions that help one's cause and underestimate the strength of those that hinder it. If one tries to compensate for such bias effects, one may be led to undercompensate or overcompensate; the standpoint of consciousness gives one no very privileged access to whether one has succeeded, for bias does not work by purely conscious processes. Its effects are much easier to observe in others than in oneself. There may be further obstacles to knowledge of one's own intuitions too. For example, philosophers with a tin ear for natural language sometimes seem to misarticulate their own strong intuitions, using forms of words that do not express what they really want to say.

The foregoing arguments do not show that we are not often or typically in a position to know what intuitions we have. The point is just that not even facts about intuition meet the fully operational standard for evidence that was used to exclude other facts. If sceptical scenarios for the claim that there are mountains in Switzerland or that the subject in a Gettier case lacks knowledge mean that our evidence in philosophy cannot include the non-psychological fact that there are mountains in Switzerland or the fact that the subject in a Gettier case lacks knowledge, then the sceptical scenarios just indicated for the claim that we have an intuition that there are mountains in Switzerland or that we have an intuition that the subject in a Gettier case lacks knowledge mean that our evidence in philosophy cannot include the psychological fact that we have an intuition that there are mountains in Switzerland or the fact that we have an intuition that the subject in a Gettier case lacks knowledge. If we go down that road, we shall soon have no evidence left. Once we relinquish the hopelessly demanding operational standard for evidence, it is unclear why our evidence should not include the non-psychological fact that there are mountains in Switzerland or the fact that the subject in a Gettier case lacks knowledge.

The retreat to confining evidence to psychological facts might be defended on more pragmatic grounds. In debate, one cannot hope to persuade opponents by appealing to evidence that they do not accept. Predictably, they will accuse one of begging the question. A fact can function as evidence in the debate only if both sides are willing to accept it. If one party asserts that P while the other party denies that P, they cannot use the fact that P as shared evidence, but they can use the fact that the first party asserts that P as shared evidence, because they presumably agree on that. More interestingly, the second party may also

feel some inclination to judge that P, but resist it because they take it to be defeated by countervailing theoretical considerations. In that case, both parties may agree that they share a conscious inclination to judge that P, an intersubjective rather than merely idiosyncratic phenomenon. They can therefore use the fact that there is a shared intuition that P as shared evidence, of which they will offer rival explanations. Perhaps the first party will say that we have the intuition that P because we recognize the obvious fact that P, while the second party says that we have the intuition that P because we are socially conditioned to do so when we are inducted into a conceptual practice that is convenient for everyday purposes but fails to fit the underlying facts and may even be incoherent. On the pragmatic view, what allows psychological facts about our intuitions to serve as evidence in a given context is that they happen to be uncontroversial in that context, not that they are uncontroversial in all contexts, or foundational in some deeper sense. Currently undisputed non-psychological truths can be used as evidence too. We can get by with agreement on particular pieces of evidence without having a fully operational standard for evidence in general. This dialectical conception of evidence is applicable even to an isolated thinker, for in isolation one can still play rival theories against each other in one's head. Virtual opponents suffice for much philosophical thinking.

The dialectical standard does not favour the use of psychological facts about our intuitions as evidence in all philosophical contexts for, as we have seen, such facts can also be controversial. Disagreements sometimes arise as to the strength of various intuitions, their genuineness or their proper expression. Some possible causes of such disagreement were noted above. Furthermore, the very idea of intuition is controversial in some philosophical circles. Radical eliminativists about the mind may say 'Research in neurophysiology has shown that folk psychology is a false theory; its ascriptions of mental states and acts are never strictly and literally true, however convenient they may have been' (let us not ask whether radical eliminativists believe what they say). On their view, humans are incapable of judging that P, because judging is a folk-psychological propositional attitude; thus humans are never inclined to judge that P, consciously or otherwise, and so never have the intuition that P. In particular, consistent radical eliminativists will not even concede that radical eliminativism is counterintuitive, or that we have the intuition that we have beliefs and desires. If one wants to find common ground with radical eliminativists, then one must rigorously depsychologize one's evidence, and replace intuitions by the results of brain scans. Similarly, sceptics about dispositions doubt that anything or anyone is ever disposed or inclined to be or do anything.

In particular, they will refuse to admit that anyone is really inclined to judge anything, and therefore refuse to admit that anyone really has intuitions. Thus the dialectical standard excludes the use of psychological facts about our intuitions as evidence in some philosophical contexts.

There are general grounds for dissatisfaction with the dialectical standard of evidence. To test one's beliefs by one's ability to persuade a sceptic of their truth is to play a dangerous game. With that narrow common ground as one's basis, one has little chance of success. If one uses only premises and forms of inference that a sceptic about perception will allow one, and therefore only premises that are true and forms of inference that are valid even if one is a brain in a vat, then one has little prospect of mounting a good argument to the conclusion that one has hands. But it is widely, although not universally, acknowledged that it does not follow that we do not know that we have hands. To be genuine, knowledge need not be recoverable from an impoverished sceptical starting-point. After all, if one uses only premises and forms of inference that sceptics about reason will allow one, then one cannot mount a good argument to the conclusion that there are good reasons. For since such sceptics doubt that there are good reasons, they will allow one neither the proposition that there are good reasons as a premise nor any form of inference (reasoning) at all with which to reach it as a conclusion from some other starting-point. But it would be frivolous to conclude, from that trivial point, that we do not know that there are good reasons. Indeed, even sceptics about reason must deny that conclusion to follow, for they deny that any conclusion follows. Sometimes, in self-defence, one must abandon sceptics to their fate. Some scepticism, such as scepticism about reason, is so radical that it leaves too little unchallenged for what remains as shared evidence to be an appropriate basis for evaluating the propositions under challenge. When one is warranted in refusing to play the sceptic's dialectical game, the dialectical standard of evidence becomes irrelevant. In so refusing, one does not abandon one's claims to knowledge and reason, because the appropriate standard of evidence is non-dialectical. By that standard, the sceptic's peremptory challenge fails to disqualify the challenged fact as evidence. One continues to assert propositions of the disputed kind on the basis of appropriate evidence, without expecting to be able to mount arguments for them using only premises and forms of inference that sceptics about propositions of that kind will allow one. Since escape from the radical sceptical predicament is impossible, one takes good care not to get into that predicament in the first place.

Is the attitude just sketched a legitimate response to scepticism about judgement? For example, may one take one's evidence to include the fact that the

subject in a Gettier case lacks knowledge or that there are mountains in Switzerland, even though the sceptic about judgement denies one the right to such evidence? In reaching one's views, one does not restrict oneself to premises and forms of inference that sceptics about judgement will allow one, for one regards their restricted evidence base as too wilfully impoverished to constitute a reasonable starting-point for evaluating the propositions in which one is interested. Sceptics about judgement have not shown that the facts that they allow as evidence are really more certain than the facts that they disallow. In particular, it is quite insufficient for them to point out that it is logically possible for us to judge that there are mountains in Switzerland even if there are no mountains in Switzerland, for it is equally logically possible for us to judge that we have the intuition that there are mountains in Switzerland even if we lack the intuition that there are mountains in Switzerland.

Even if (let us pretend) facts about our intuitions were in some sense more certain for us than all other facts, it would not follow that we should restrict our evidence base to facts about our intuitions. For the extra information in a much wider evidence base might be worth the cost in a slight loss of reliability. After all, if logical truths were more certain than all other facts, it would not follow that we should restrict our evidence base to logical truths: that would prevent us from gaining any empirical knowledge. It would be scepticism about everything except reason.

One might find this short way with the sceptic about judgement contrary to the open spirit of philosophical discussion. The sceptic has serious concerns of a recognizably philosophical kind: do they not deserve a fair hearing? How can they be given such a hearing if the very propositions that the sceptic challenges are taken as evidence? Sceptics of any principled kind can indeed expect more tolerance in philosophy than in other disciplines. One can discuss their scepticism with them without stepping outside the bounds of philosophy. In talking to them, it is pointless to offer for their acceptance propositions that one knows them to be unwilling to accept. In particular, it seems unphilosophical to refuse to discuss scepticism about judgement with its proponents. In conversation with them, it is dialectically pointless, rude, to offer as evidence propositions that one knows them to reject. But the issue remains: what implications, if any, does the outcome of such a conversation have for the epistemic status of belief in the propositions that the sceptic questions? Faced with a sceptic about reason, or everything except reason, many philosophers would be willing to start a conversation, out of politeness, curiosity, competitiveness or the desire to save a soul. But their inability to achieve a dialectical triumph over such a resourceful opponent does not oblige them to become sceptics

about reason, or everything except reason, themselves. There is no bad faith in continuing to claim knowledge of the contested truths. For the anti-sceptic is not obliged to treat dialectic as the measure of all things. Indeed, the claim that dialectic is the measure of all things faces self-defeat, for it cannot triumph dialectically over its denial; even if it appeared to be getting the better of the argument, would not taking that to establish its truth beg the question? Similarly, even if one cannot establish dialectically, in dispute with the sceptic about judgement, that the subject in a Gettier case lacks knowledge or that there are mountains in Switzerland, without bad faith one can still claim to know that the subject in a Gettier case lacks knowledge or that there are mountains in Switzerland.

The conception of our evidence in philosophy as merely psychological facts to the effect that we have various intuitions does not withstand epistemological scrutiny. But should we think of such facts concerning our own intuitions as merely psychological facts about us anyway?

3. Intuition and conservativism.

According to scepticism about judgement, that we have the intuition that P is a merely psychological fact about us: typically, it has no bearing on the question whether P. That attitude is hard to sustain with respect to one's own intuitions. To have the intuition that P is to be consciously inclined to judge that P. Suppose that I judge according to my inclinations. Judging that P is the active form of believing that P. How can I regard it as a merely psychological fact about myself that I believe that P? That P entails that whoever believes that P has a true belief that P, for the belief that P is true if and only if P. Therefore, in believing that P, I am committed to the belief that the belief that P is true, and to the belief that I must continue believing that P if I am to continue believing the truth as to whether P.[7] Thus I cannot be cognitively neutral towards my own beliefs. I must regard them as true beliefs, not as mere psychological phenomena. Consequently, when I judge as I am inclined to, I cannot be cognitively neutral towards my own intuitions. I am committed to them in a way in which I am not committed to my own fantasies. This commitment might be thought to explain the epistemic role of intuitions. They are what we start from, the boat that we find ourselves in. Perhaps we can progressively replace them, but we cannot distance ourselves from all of them at once, for we have nowhere else to stand. That is why one is bound to give primacy to one's

---

[7]   For simplicity, propositions are assumed to be so individuated that they cannot change in truth-value over time.

own intuitions over those of other people in one's philosophical thinking, even when one knows only too well that the others have conflicting intuitions. By contrast, when one treats intuitions as mere psychological phenomena, in the evaluation of a psychological theory, data about one's own intuitions carry no greater evidential weight than do data about the intuitions of others: but when one judges according to one's inclinations, one cannot treat one's own intuitions in that psychologistic manner.

The practical necessity of starting from where one is may be elevated to normative status in a principle of epistemic conservativism: that one has a defeasible right to one's beliefs, which right may be defeated by positive reasons for doubt, but not by the mere absence of independent justification.[8] Thus one's belief that the subject in a Gettier case lacks knowledge or that there are mountains in Switzerland gives one the defeasible right to rest arguments on the premise that the subject in a Gettier case lacks knowledge or that there are mountains in Switzerland.

In order to assess such an account of the role of intuitions, we must be clearer about the relation between intuition and belief. For David Lewis, 'Our "intuitions" are simply opinions' (Lewis 1983a, x). Are intuitions just beliefs? In the previous sense, an intuition that P is a conscious inclination to judge that P. For Peter van Inwagen, 'Our "intuitions" are simply our beliefs – or perhaps, in some cases, the tendencies that make certain beliefs attractive to us, that "move" us in the direction of accepting certain propositions without taking us all the way to acceptance' (van Inwagen 1997, 309; he adds parenthetically 'Philosophers call their philosophical beliefs intuitions because "intuition" sounds more authoritative than "belief"'). Even if one discounts unconscious beliefs and the difference between the state of believing and the act of judging, one can still distinguish the two accounts of intuition. For we do not always follow our inclinations; someone inclined to believe that P may nevertheless not believe that P. I am tempted to believe that a smooth stone is smooth even at the microscopic level, but I resist the temptation because I know better. In this respect, the analogy between intuitions and perceptual appearances is apt. When the stick looks bent in water, I am tempted to believe that it is bent, but I resist the temptation because I know better. Background

---

8   See Harman (1986, 29-42) for a defence of epistemic conservativism, and Vahid 2003 for a recent critical survey of its varieties. For the sake of both simplicity and generality, subtleties in the formulation of the principle have been glossed over. The closely related method of reflective equilibrium described by Goodman 1965 and Rawls 1971 shows how a philosophical position can evolve from an original set of intuitions; it is discussed in several essays in DePaul and Ramsey 1998.

information can defeat our inclination to take perceptual or intellectual appearances at face value. The difference between belief and inclination to belief may matter for the principle of epistemic conservativism.

Suppose that Justin has been brought up to believe that knowledge is to be analysed as justified true belief. He is confronted for the first time with a Gettier case. He might immediately judge with confidence that the subject has justified true belief without knowledge, and abandon his old belief that justified true belief is necessary and sufficient for knowledge; presumably, epistemic conservativism then switches sides and starts supporting his new belief that justified true belief is not necessary and sufficient for knowledge. But suppose instead that Justin is more cautious, not wanting to assent to anything tricky too readily. Although he is consciously inclined to judge that the subject has justified true belief without knowledge, he does not immediately give in to that inclination, nor does he immediately abandon his ingrained belief that justified true belief is necessary and sufficient for knowledge. It is not clear that epistemic conservativism counsels abandoning the traditional analysis in this situation. Although the Gettier case might be counted a positive reason for doubting the traditional analysis, epistemic conservativism as formulated above does not explain why it should. If Justin is asked 'What reason is there to doubt the traditional analysis?', the answer 'The subject in this possible case has justified true belief without knowledge' would obviously be relevant, since it expresses a proposition inconsistent with the traditional analysis. But he is in a position to give that answer only if he already believes that the subject in the Gettier case has justified true belief without knowledge. Since he does not already believe that, he has to say something else. The answer 'I am inclined to believe that the subject in this possible case has justified true belief without knowledge' might be relevant if the function of the words 'I am inclined to believe that' were to signal tentative assent to the proposition expressed by what follows, but Justin's commitment to the traditional analysis may disincline him to give even tentative assent to a putative counterexample unless he is forced to do so. If the function of the words 'I am inclined to believe that' is instead to report his psychological state of being inclined to believe the proposition expressed by what follows, as their literal meaning suggests, then the relevance of the answer to the original question is far from obvious, for he has not yet given even tentative assent to the Gettier case as a counterexample.

Might epistemic conservativism be extended to claim that one has a defeasible right to believe whatever one is inclined to believe? Such an extension is much less clearly motivated than the original principle by the idea that,

since one must start from where one is, one has at least a defeasible right to be where one is. A right to be where I am is a right to have the beliefs and inclinations that I have. That does not obviously include a right to follow those inclinations to new places, especially when the beliefs that I already have imply that those are bad places to go to, for example, when the inclinations are to believe things inconsistent with my current beliefs. As the Gettier counterexamples show, intuition can be revolutionary as well as conservative. If I currently believe that P, then I am currently committed to the belief that any inclination to believe something that entails that it is not the case that P is an inclination to believe something false. In the way in which I am committed to the propositions that I believe, I am not committed to the propositions that I am merely inclined to believe; I am merely inclined to commit myself to them in that way. After all, a right to be where I am is of limited practical use unless it involves a right to stay where I am, to continue believing, at least for a while, what I currently believe.

A further problem faces the attempt to base a response to scepticism about judgement on the first person present tense of 'believe'. Even if it explains why one cannot regard one's own current beliefs as the products of a sceptical scenario, it does not explain why one should not think that most beliefs are such products. If I believe that P, then I am committed to the belief that the belief that P is true. That point can be generalized over all subjects at all times in all possible worlds. But, obviously, the correct generalization does not imply that if I believe that P, then you are committed to the belief that the belief that P is true, for you may consistently believe that I am wrong – although of course if you do believe that P, then you are committed to the belief that the belief that P is true. Similarly, the correct generalization does not imply that if I believed yesterday or shall believe tomorrow that P, then I am committed today to the belief that the belief that P is true; I may believe today that my evidence yesterday or tomorrow is misleading – although of course if I believed yesterday or shall believe tomorrow that P then I was committed yesterday or shall be committed tomorrow to the belief that the belief that P is true (see also Williamson 2000, 219). Nor does the correct generalization imply that, if I would have believed that P if things had been slightly different, then I am actually committed to the belief that in those counterfactual circumstances the belief that P would have been true – although of course if I believed in those circumstances that P then I would be committed in those circumstances to the belief that the belief that P was true. While strongly committed to the truth of my own actual present beliefs, I might regard the beliefs of other subjects and my own beliefs at other times or in other possible circumstances as in massive

error. I might take sceptical scenarios to prevail almost everywhere while insisting that I happen not to be currently in one. But such a response to scepticism would be unimpressive, perhaps unstable. The admitted frequency of sceptical scenarios in nearby situations constitutes an urgent reason for doubting one's own beliefs; it triggers the defeating condition in the principle of epistemic conservativism. One should beware of regarding oneself as too happy an exception to sadly general trends.

Few of us regard ourselves as highly exceptional in having escaped the worst scenarios for scepticism about perception. We think that such scenarios are rare in worlds like ours. We find the brain in a vat scenario far-fetched. The environment as we perceive it is full of creatures in regular perceptual contact with that very environment. It takes no special luck or skill to avoid envatment: human beings have never been in danger of it. Of course, sceptics will say that such empirical claims about our environment merely beg the question against them, for their truth is part of what the sceptic questions. But the empirical claims were not addressed to sceptics, in a futile attempt to persuade them out of scepticism. Instead, they figure in our appraisal of sceptical arguments, from our current non-sceptical point of view.[9] We have to decide, without yet having suspended our ordinary empirical beliefs, whether the acknowledged bare metaphysical possibility of the sceptical scenarios gives us good reason to suspend those empirical beliefs – not just for a few minutes in an epistemology seminar, but for the rest of our lives. Most of us find the reason inadequate. Bare possibilities of error, however picturesque, do not constitute an imminent danger; the threat is not nearly urgent enough to warrant the drastic and costly precautions that the sceptic recommends. For most purposes, we do not take the sceptical possibilities seriously.

Our tendency to ignore sceptical possibilities cannot be explained by their making no practical difference, for many of them make such a difference. If you are a brain in a vat who does not really interact with other people, then much of your altruistic behaviour is futile. Furthermore, there is a sceptical scenario in which you will be subjected to unremitting horrible pain for years, starting tomorrow, unless you immediately do what appears to you exactly like going out and buying ten copies of the same newspaper: you do not even take that elementary precaution (I bet). Of course, in another sceptical scenario, you will be subjected to unremitting horrible pain for years, starting tomorrow, if you immediately do what appears to you exactly like going out and buying ten copies of the same newspaper. If one takes all possibilities equally se-

---

⁹ Compare Nozick 1981, 167.

riously, they tend to cancel each other out for practical purposes. But that does not imply that we are left back where we were before sceptical possibilities occurred to us. If everything except your present consciousness is totally unknown, why not simply indulge in sweet dreams? According to the thoroughgoing sceptic, it is no more probable (epistemically) that you have hands than that you are in a sceptical scenario in which you merely appear to have hands: will you therefore reject a bet on which you win 10 euros if you have hands and lose 100 euros otherwise on the grounds that its expected utility is negative, since 10/2 - 100/2 = -45 (if scepticism makes you doubt the enforceability of the bet, that is no reason to accept it)? Surely it is a good bet, even if you happen to be in an epistemology seminar at the time. We ignore radical sceptical possibilities in practice, even when they are drawn to our attention, because we do not rate them as epistemically serious possibilities. We make that epistemic assessment from our non-sceptical perspective.

When we judge that in our world radical sceptical scenarios present no imminent danger to anyone, we do so on the basis of our own beliefs, but that judgement depends on the specific content of those beliefs; it is not an automatic consequence of epistemic conservativism by itself. We have a rich conception of ourselves and our environment on which brains in vats are very far-out physical possibilities. That conception also enables us to fill in specific answers to the question 'How do you know?' as it arises on specific occasions, for example by indicating relevant processes of perception, memory, testimony and inference, although of course the conception need not figure amongst premises from which the more specific knowledge was inferred, since the latter need not have been inferred at all. None of this amounts to a detailed dissection of the flaws in particular sceptical arguments. Rather, it provides the appropriate background to our confidence that such flaws must be there.

Can we respond in the way just sketched to scepticism about judgement? On our conception of our world, brains in vats are distant possibilities: what of the scenarios on which the sceptic about judgement relies? Scepticism about perception starts with actual perceptual errors and imaginatively radicalizes them until it reaches brains in vats. Similarly, scepticism about judgement starts with actual errors about witchcraft, oracles and magic and imaginatively radicalizes them until it reaches the nonexistence of mountains. In both cases, there is a trade-off between how remote the sceptical scenarios are (judged from our current perspective) and how far-reaching a scepticism they motivate. Very close possibilities motivate only a very limited scepticism; more remote possibilities motivate a more general scepticism. The closer the possibility, the more seriously it deserves to be taken. For scepticism about per-

ception, we have at least a rough idea of what makes the more radical scenarios remote, the enormous practical obstacles to setting up all the requisite causal mechanisms, not to mention the shortage of motivation for doing so. For scepticism about judgement, what corresponds to those obstacles? Do we even believe that the actual world is not full of apt scenarios for such scepticism?

Suppose that most ordinary beliefs in most other cultures are false, because the constituent concepts of those beliefs are somehow laden with false theories.[10] Then the possibility that, for similar reasons, most ordinary beliefs in our own culture are also false is too close to home to be dismissed as fanciful or far-fetched. Scepticism about judgement gets a grip. A satisfying response would put such sceptical scenarios far from other cultures, not just far from one's own.

Given empirical evidence for the approximate intertranslatability of all human languages and a universal innate basis of human cognition, we may wonder how 'other' any human culture really is. If we believe that P and that others believe that P too, then we are committed to the belief that the others' belief that P is true. But if the tendency of human beliefs to be true is merely an accidental byproduct of our DNA, and other galaxies are rife with non-human persons most of whose beliefs are false, because their constituent concepts are laden with false theories, then scenarios for scepticism about judgement are still dangerously close to home. Even if such scenarios are rare or absent in the actual universe, but only by good luck, it remains uncomfortable for the opponent of scepticism about judgement. If we are to refuse in good conscience to take seriously the radical scenarios for scepticism about judgement, we must do so from a perspective on which there is a quite general tendency for beliefs to be true. Anything less than that will look like special pleading on our own behalf.

4. Intuition and charity.

Some naturalists will argue on evolutionary grounds that beliefs tend to be true, for creatures with too many false beliefs are unfit to survive. True beliefs tend to cause one to get what one wants in a way in which false beliefs do not. Truth conduces to success. That is not to deny that some false beliefs have survival value; the suggestion is only that on the whole truth is more conducive than falsity to survival. Since we are arguing from our current perspective, on which our world is governed by regularities that extend over past, present and

---

[10] For present purposes it matters little how coarsely or finely cultures are individuated.

future, we need not worry overmuch about scenarios for inductive scepticism on which generalizations with only true instances up to some future time *t* have false instances thereafter (in any case, scepticism about judgement is not scepticism about induction). We can take past success as some guide to future success.

How do true beliefs tend to cause success in action? This principle seems central to the nature of belief and desire:

(1)   If an agent desires that P, and believes that if it does A then P, then *ceteris paribus* it does A.

The '*ceteris paribus*' clause in (1) covers the rationality of the agent, the absence of countervailing desires, and so on. Now if the agent desires that P, believes that if it does A then P, and does A, then P if the belief is true, so its desire is realized. If its belief is not true, then it may well not happen that P. Of course, that P may not help the agent if it is not good for the agent that P. We might therefore construe the argument as supporting a stronger conclusion: that evolution favours creatures who both believe what is true and desire what is good for them.[11]

The act A is not something of which the creature has no idea. It conceives A in believing that if it does A then P. If an agent does A without believing itself to be doing so, then the natural link between antecedent and consequent in (1) is broken. For example, if you go north while believing that you are going south, your action is not explained just by your desire to reach the oasis and belief that if you go north then you will reach the oasis, (1) notwithstanding. Perhaps the explanation is that, in addition, you desire even more strongly to avoid your enemy and believe that he is at the oasis. Although such examples do not refute (1), since the '*ceteris paribus*' clause absorbs their shock, they indicate that the rationale for (1) takes for granted that beliefs about what one is doing tend to be true, which is a special case of the very phenomenon that we are trying to understand. Therefore, in order not to assume what needs to be explained, let us revise (1) thus:

(2)   If an agent desires that P, and believes that if it does A then P, then *ceteris paribus* it acts so that it believes that it does A.

---

[11] If desiring that P were believing that it is good for one that P, the tendency to desire what is good for one might be subsumed under the tendency to believe what is true: but one can desire that P because one believes that it is good for everyone else, but not for oneself, that P; altruism is possible.

Given that you want to avoid your enemy, and believe that if you go south then you will avoid him, one can use (2) to explain why you act so that you believe that you go south, even though in fact you go north. Moreover, starting from (2) rather than (1), one can still explain why it is good for an agent to have true beliefs and desires for what is good for it. For if it desires that P, believes that if it does A then P, and acts so that it believes that it does A, then P if both beliefs are true, which is good for it if its desire is for what is good for it.

Unfortunately, such a derivation explains much less than it appears to. For, given (2), one can show in the same way for infinitely many deviant properties true* and good* that the combination of true* beliefs and desires for what is good* for one yields (*ceteris paribus*) what is good (not just good*) for one. To see this, consider an arbitrary mapping on propositions, taking the proposition that P to the proposition that ^P, subject to the constraint that it commutes with logical operations, in the sense that the proposition that ^(it is not the case that P) is the proposition that it is not the case that ^P, the proposition that ^(if P then Q) is the proposition that if ^P then ^Q, and so on. In other respects, the mapping is arbitrary: for example, the proposition that ^(I am going north) might be the proposition that you are eating slowly.

If a proposition is just the set of possible worlds in which it is true, then we can construct such a mapping for any permutation π of possible worlds (a one-one mapping of the possible worlds onto the possible worlds) by stipulating that each world w belongs to the proposition that ^P if and only if $\pi(w)$ belongs to the proposition that P. The mapping commutes with negation, for example, because, for any world w, the following are equivalent: w belongs to the proposition that ^(it is not the case that P); $\pi(w)$ belongs to the proposition that it is not the case that P; $\pi(w)$ does not belong to the proposition that P; w does not belong to the proposition that ^P; w belongs to the proposition that it is not the case that ^P. For similar reasons the mapping commutes with other logical operations, such as the truth-functional conditional.

Alternatively, if propositions have quasi-syntactic structure, then we can take an arbitrary mapping on their atomic constituents and extend it recursively to complex propositions in the natural way. The mapping automatically commutes with logical operations because the commutativity clauses are built into its inductive definition.

Now define 'true*' and 'good*' by these equivalences:

(3)   That P is true* if and only if that ^P is true.
(4)   That P is good* for an agent if and only if that ^P is good for it.

Now suppose that an agent desires that P, believes that if it does A then P, and acts so that it believes that it does A. Suppose further that both beliefs are true*. By (3), since the proposition that if it does A then P is true*, the proposition that ^(if it does A then P) is true. Since the mapping commutes with logical operations, in particular with the truth-functional conditional employed (by stipulation) in (1) and (2), the proposition that ^(if it does A then P) is the proposition that if ^(it does A) then ^P. Thus the proposition that if ^(it does A) then ^P is true. By (3) again, since the proposition that it does A is true*, the proposition that ^(it does A) is true. Since truth is closed under modus ponens, the proposition that ^P is true. Suppose finally that what the agent desires is good* for it. Thus that P is good* for it; therefore, by (4), that ^P is good for it. In other words, something (that ^P) is the case that is good for the agent: together, true* belief and desire for what is good* for one yield (*ceteris paribus*) what is good (not just good*) for one.

From (2), we cannot conclude that the combination of true belief and desire for what is good for one is any better for one than the combination of true* belief and desire for what is good* for one. Yet, despite all the evolutionary pressures, we have no special tendency to believe what is true* or to desire what is good* for us. For example, that I am going north may be true* if and only if you are eating slowly, and that I reach the oasis may be good* for me if and only if it is good for me that you read your book. I have no special tendency to believe that I am going north only if you are in fact eating slowly or to desire that I reach the oasis only if it is in fact good for me that you read your book. If we start without any correlation between belief and truth, considerations of survival will not make the connection for us.

Suppose that we are trying to understand some aliens. We have an extremely plausible interpretation Int of their beliefs and desires. We define a new interpretation Int* by specifying that, under Int*, an alien believes that ^P if and only if, under Int, it believes that P, and, under Int*, it desires that ^P if and only if, under Int, it desires that P.[12] Thus Int* ascribes a true belief just where Int ascribes a true* belief; Int* ascribes a desire for what is in fact good for one just where Int ascribes a desire for what is in fact good* for one. Int* attributes bizarre contents to the aliens: under Int*, their beliefs about their environment have no tendency to be true, their bodily movements no tendency to bring about the satisfaction of their desires. For example, under Int, an alien desires that it will be cool and believes that if it jumps into the lake then it will

---

[12] The definition of Int* assumes that the proposition that ^P is the proposition that ^Q if and only if the proposition that P is the proposition that Q; this condition is easily met. Int* is also stipulated to ascribe to the aliens only beliefs and desires of the form that ^P.

be cool; it jumps into the lake and will be cool. Under Int*, it desires that ^(it will be cool) and believes that ^(if it jumps into the lake then it will be cool), in other words, that if ^(it jumps into the lake) then ^(it will be cool); it jumps into the lake and will be cool. For definiteness, let that ^(it will be cool) and that ^(it jumps into the lake) be that you were tall and that you went to bed respectively. Thus, under Int*, the alien desires that you were tall and believes that if you went to bed then you were tall; it jumps into the lake and will be cool. Under Int, when it jumps into the lake it also believes that it jumps into the lake and that it will be cool. Thus, under Int*, when it jumps into the lake it believes that you went to bed and that you were tall. Int* make the aliens' mental lives formally as rational and coherent in propositional content as Int does; but Int* radically disconnects their mental lives from what is happening around them and from what they are physically doing, whereas Int keeps them connected in the normal way. Moreover, Int* postulates no special mechanism to help explain the strange disconnection. Surely Int* would misinterpret the aliens. Even if such radical disconnection is not metaphysically impossible, it would occur only under highly abnormal circumstances. The nature of mental content seems to favour Int over Int* in some constitutive way. What is that way?

We could try to rule out Int* by proposing more specific constraints on the internal interconnections of propositional attitudes for Int* to fail. But that approach is quite unpromising; it misses the point of the problem. There is plenty of scope for tailoring the deviant interpretation Int* to meet even the more specific internalist constraints while still attributing mental lives radically disconnected from the environment and bodily behaviour. Rather, we need constraints on the relation between mental life and the external world. Much contemporary philosophy consists of attempts to provide such constraints.

The attempts may be divided into the molecular and the holistic.[13] The molecularist analyses mental contents into constituents, and for each constituent tries to specify conditions for thinking with it. For example, a simple theory of possession conditions for concepts says that to possess the concept *mountain* one must, under optimal conditions specified without ascription of that very concept, be willing to judge *here is a mountain* if and only if a mountain is present. A simple verificationist theory of meaning states necessary and sufficient conditions for the sentence 'Here is a mountain' to be canonically verified (or assertible). A simple causal theory of reference says that a thought token refers to mountains if and only if it is causally related in a specified way

---

[13] The terminology of 'holism' and 'molecularism' is hijacked from Dummett 1975 to make a slightly different distinction.

to mountains. And so on. More complex and sophisticated accounts can be developed in the same spirit.

If a molecularist account could be made to work, it might support many of the conclusions of this paper. However, molecularist accounts face major obstacles. For instance, it is hard for an account that is intended to provide non-circular necessary conditions for concept possession to say anything non-trivial about what the subject does in non-optimal conditions, where ignorance and error are rife even amongst those who possess the concepts at issue; yet it is hard for an account intended to provide non-circular sufficient conditions for concept possession to say nothing non-trivial about what the subject does in non-optimal conditions.

It is also hard to screen out the effects of the subject's background theory without circularity. For example, if the optimal conditions are specified without ascription of the concept *mountain*, then they can presumably be met when a revisionary metaphysician, a native English speaker with good eyesight and open eyes, dissents in good visibility from the sentence 'Here is a mountain' in the middle of the Alps. The danger is that the possession condition would count her as lacking the concept *mountain,* a highly implausible result. By any reasonable standard she had the concept *mountain* before she developed her revisionary metaphysics; since she fully understood the English word 'mountain', she knew that it means *mountain*. Developing her revisionary metaphysics did not make her cease to understand the word 'mountain'; she understands the word in the normal way as used by other speakers, and therefore knows that it means *mountain*; she still has the concept *mountain*. When she denies that there are mountains, she is consciously disagreeing with common sense, not talking past it. Similar problems plague verificationist theories of meaning and even causal theories of reference.[14]

It is similarly unclear how the causal connections are supposed to work in non-optimal conditions or for speakers with non-standard background theories. Nor are causal connections always needed. Even for mountains, a community might think about them without ever having had any causal contact with them, by having causal contact with hills and thinking of mountains as like hills, only bigger.

The history of molecularist programmes gives no grounds for optimism that such obstacles will eventually be overcome. That is not to imply that all molecularist claims are hopelessly false. Many of them seem to be true 'for

---

[14] This argument is further developed in Williamson 2003. Although the arguments of Quine 1951 against traditional conceptions of meaning and the analytic-synthetic distinction look much weaker than they once did, his Duhemian challenge to the idea of sentences whose meaning compels assent from one independently of one's background theory remains apt.

the most part'. What is doubtful is that they can be replaced by strictly true claims that remain within the spirit of a molecularist programme.

The alternative to molecularism is holism. Although holism need not deny that thoughts have constituent structure, its constraints on thinking given thoughts apply at the level of the subject's total system of thoughts, not at the level of individual constituents; they are global rather than local. The most salient holistic proposal is Donald Davidson's principle of charity. According to Davidson, 'Charity is forced on us; whether we like it or not, if we want to understand others, we must count them right in most matters' (Davidson 1974, 197). He argues that good interpretation imputes agreement in the main between interpreter and interpreted; we can coherently conceive an omniscient interpreter, agreement with whom guarantees truth; since the omniscient interpreter's interpretation is by hypothesis correct, good interpretation imputes truth in the main (Davidson 1977, 200-1). Thus, by Davidson's lights, the revisionary metaphysicians are bad interpreters if they interpret ordinary people as in massive error, for example over the existence of mountains. Of course, a revisionary metaphysician might claim that ordinary people do not really believe that there are mountains, but that seems to be an even worse misinterpretation. Davidson's account directly implies a tendency for beliefs to be true.

Davidson's principle of charity evokes massive disagreement. However, it is not wholly to blame for the contentious conclusions that Davidson uses it to draw. It figures in his notorious argument against the very idea of mutually incommensurable conceptual schemes, alien ways of thought or untranslatable languages (Davidson 1974). But that argument also makes both the verificationist assumption that other creatures have beliefs only if we can have good evidence that they have beliefs and the constructivist assumption that we can have good evidence that they have beliefs only if we can have good evidence as to which beliefs they have. Neither assumption follows from the principle that beliefs tend to be true. Neither assumption is warranted, for we are far from omniscient interpreters (compare Nagel 1986, 93-99). The aliens may be able to interpret each other even if we cannot interpret them. More generally, Davidson's application of the methodology of radical interpretation to the philosophy of language embodies a kind of ideal verificationism, on which agents have just the intentional states that a good interpreter with unlimited access to non-intentional data would ascribe to them. However, we could, as David Lewis (1974, 110-11) recommends, treat the predicament of the radical interpreter as merely a literary device for dramatizing the question: how do the intentional states of agents supervene on the non-intentional states of the world? The sense in which that question concerns the determination of content is

metaphysical, not epistemological. In this spirit, we could consistently accept a principle of charity while allowing that alternative conceptual schemes are possible.[15]

If the role of the radical interpreter is inessential, so too is that of agreement between interpreter and interpreted. Truth has priority over agreement: the metaphysical version of Davidson's principle of charity requires that agents have mostly true beliefs. That is in effect a constraint on reference for the constituents of beliefs or of the sentences that express them. Agreement is secondary; two agents with mostly true beliefs do not mostly disagree with each other, although they may have few beliefs in common, if they have different concerns, and may even tend to disagree over their limited common concerns.

Davidson's principle of charity is too loose to figure in an algorithm for reducing the intentional to the non-intentional. But present purposes do not require us to engage in the heroically ambitious quest for such a reduction. What we need are correct nontrivial principles about propositional attitudes that somehow link belief and truth, metaphysically rather than epistemologically. Such principles can fall far short of reducing the intentional to the non-intentional, even of fixing the supervenience of the former on the latter.

Even in its de-epistemologized, non-reductive version, Davidson's principle of charity remains highly contentious. Massive error seems genuinely possible for a brain envatted only months ago.[16] Philosophers have responded by formulating revised principles that allow one to interpret another as in massive error when one would have been in massive error oneself in her circumstances. For example, Richard Grandy proposes 'as a pragmatic constraint on translation' a *principle of humanity*: 'the condition that the imputed pattern of relations among beliefs, desires, and the world be as similar to our own as possible' (Grandy 1973, 443). Even if we treat the principle of humanity as a metaphysical constraint on content, rather than a pragmatic constraint on translation, it says nothing directly about any tendency for beliefs to be true. However, since each of our beliefs commits us to regarding it as true, and therefore as having that relation to the world, one could argue that the principle of humanity requires the beliefs of others to tend to have the same relation to the world, and therefore to be true too. Perhaps humanity implies at least a limited ver-

---

[15] By contrast, McGinn treats radical interpretation as an epistemological problem, and explicitly allows for uninterpretable believers (McGinn 1986, 367). For a recent discussion of Davidson on radical interpretation, see McCulloch 2003, 94-108. For simplicity, the present paper ignores several vast issues about radical interpretation (the relation between thought and language; indeterminacy and inscrutability; …).

[16] Klein 1986 discusses of Davidson's treatment of sceptical scenarios. See McCulloch 2003, 126-40 for a recent discussion of the difficulty of interpreting brains in vats.

sion of charity, although the vagueness of 'similarity' between patterns of relations makes it hard to tell. But the anthropocentrism of the principle of humanity is suspect. After all, we humans are prone to peculiar logical and statistical fallacies: once we recognize a quirky design fault in ourselves, it would be perverse to prefer, on metaphysical principle, interpretations of non-human aliens that attribute the same design fault to them. Although humans are the clearest examples of rational agents with which we are familiar, we are also clear that there could be far more rational agents than we are. On their metaphysical reading, anthropocentric principles of charity implausibly imply that the very nature of content militates against the possibility of superhuman rationality.

Other principles of charity put a premium on rationality or coherence, internalistically conceived. But they do not explain the superiority of the sensible interpretation Int over the silly Int* above. Even those that enjoin the minimization of *inexplicable* error or ignorance rely on there being further principles, so far unspecified, for explaining error and ignorance when they are legitimately attributed: whatever those further principles are, they will do much of the work in specifying the relations between mind and world. We need to make a new start.

5. *Knowledge maximization.*

Suppose that Emanuel has an ill-founded faith in his ability to discern character and life-history in a face. On that basis he forms elaborate beliefs about passers-by, in which he is confident enough to bet large sums when the opportunity offers, which it rarely does. By sheer luck he has won such bets so far, which has increased his confidence in his powers, although many other beliefs that he has formed in this way are in fact false. Now Emanuel sees a stranger, Celia, standing some distance away. Looking at her face, he judges 'She is F, G, H, …'; he ascribes a character and life-history in considerable detail. In fact, none of it fits Celia. By pure coincidence, all of it fits someone else, Elsie, whom Emanuel has never seen or heard of. Does the pronoun 'she' as used by Emanuel in this context refer to Celia or to Elsie? Which of them does he use it to express beliefs about? He accepts 'She is standing in front of me', which is true if 'she' refers to Celia but false if it refers to Elsie. However, he also accepts 'She is F', 'She is G', 'She is H', …, all of which are false if 'she' refers to Celia but true if it refers to Elsie. We may assume that the latter group far outweighs the former. A principle of charity that crudely maximizes true belief or minimizes error therefore favours Elsie over Celia as the referent of the pronoun in that context. But that is a descriptive theory of

reference gone mad. Emanuel has no beliefs about Elsie. He has many beliefs about Celia, most of them false. In virtue of what is Emanuel thinking about Celia rather than Elsie?

A causal theorist of reference will point out that Emanuel's use of 'she' in this context is causally related to Celia. Of course, it may be causally related to Elsie too – she may have saved Celia's life by performing the plastic surgery on Celia's face that helped cause Emanuel's beliefs – but not in the right way for reference, whatever that is. In this case, the specific link is that Emanuel is looking at Celia and using 'she' as a visual demonstrative. But to say that he is using 'she' as a visual demonstrative is to say little more than that he is using it so as to refer to what he is looking at, and we may hope to say something more useful about what sets up this link between vision and reference. A natural idea is this. The perceptual link from Celia to Emanuel matters because it is a channel for knowledge. If 'she' refers to Celia, then, in the circumstances, Emanuel expresses knowledge when he says 'She is standing in front of me', although of course not when he says 'She is F', 'She is G', 'She is H', …, since they are false. If 'she' refers to Elsie, then of course Emanuel does not express knowledge when he says 'She is standing in front of me', since it is false, but he also fails to express knowledge when he says 'She is F', 'She is G', 'She is H', …, even though they are true. Emanuel is in a position to know of Celia that she is standing in front of him; he is not in a position to know of Elsie that she is F, G, H, …. The same contrast holds, more fundamentally, at the level of thought. The assignment of Elsie as the referent in Emanuel's beliefs gains no credit from making them true because it does not make them knowledge. The assignment of Celia wins because it does better with respect to knowledge, even though it does worse with respect to true belief.

Such examples are of course just the analogue for demonstrative pronouns of examples that Kripke and Putnam used to refute descriptive theories of reference for proper names and natural kind terms. In effect, such theories are special cases of a truth-maximizing principle of charity. Perhaps the fundamental mistake in descriptive theories of reference is to try to make true belief do the work of knowledge.

As for causal theories of reference, the postulated link between knowledge and reference suggests a schematic explanation of both their successes and their failures. Roughly: a causal connection to an object is a channel for reference to it if and only if it is a channel for the acquisition of knowledge about the object. Often, a causal connection is a channel for both. Equally, a non-causal connection to an object is a channel for reference to it if and only if it

is a channel for the acquisition of knowledge about the object. Sometimes, a non-causal connection is a channel for both. It was in any case clear that causal theories of reference and causal theories of knowledge were closely linked in their successes and failures. Both faced the problem of deviant causal chains, of specifying which causal chains carry the relevant intentional link. Both faced the problem of mathematics, which appears to exhibit both non-causal reference to abstract objects and non-causal knowledge about them.

The proposal is to replace true belief by knowledge in a principle of charity constitutive of content. But how can doing so help with the objection that massive error is possible? Presumably knowledge implies true belief. Unless the agent is inconsistent, any case of massive error is also a case of massive ignorance. At first sight, the objection only makes the problem worse. However, it is independently obvious that our knowledge is dwarfed by our ignorance. The right charitable injunction for an assignment of reference is to maximize knowledge, not to minimize ignorance (which is always infinite).[17]

Suppose that under some assignment of reference a brain in a vat has mainly true beliefs about electrical impulses in the computer that controls it. If we are still disinclined to accept the assignment, a natural reason to give is that the brain is not in a position to know about the electrical impulses. If we are inclined to accept the assignment, we are likely to think that the brain is in a position to know about them.

Here is a simpler case. A fair coin has been tossed. In fact it landed heads. The agent cannot see or otherwise know which way up it landed, but is easily convinced by what are really just his own guesses. He sincerely asserts 'Toda'. Is a point in favour of interpreting 'Toda' to mean 'It landed heads' rather than 'It landed tails' that it has him speaking and believing truly rather than falsely? Surely not. The true belief would no more be knowledge than the false belief would be. Although Davidson's principle of charity does not imply that 'Toda' cannot mean 'It landed tails', since data from other cases might outweigh the current data, it does imply that this case provides a defeasible consideration in favour of interpreting 'Toda' as 'It landed heads' rather than 'It landed tails', which it does not. The point extends to less irrational beliefs. If we interpret someone as judging on purely probabilistic grounds that ticket n did not win the lottery, our interpretation gains or loses no credit dependent on whether

---

[17] The substitution of knowledge for truth in a principle of charity is proposed in connection with a knowledge-based account of assertion at Williamson 2000, 267.

ticket *n* did in fact win, since either way the agent in the circumstances could not have known that it did not win.[18]

As already emphasized, knowledge maximization is to be understood as primarily a metaphysical principle about the nature of reference, only secondarily as an epistemological principle about the proper methodology for the ascription of reference, insofar as the former determines the latter. The nature of reference is to serve knowledge. The idea is obviously and regrettably vague; but it is not vacuous, as we have already seen.

Knowledge maximization is not in danger of absurdly imputing knowledge of quantum mechanics to stone age people. For they were not in a position to know about quantum mechanics, so assigning quantum mechanical properties or relations as the referents of their terms would not yield an interpretation on which they knew about those properties or relations. Objective limits on what subjects are in a position to know appropriately constrain the maximization of knowledge by the assignment of reference. Unless it is raining, one does not know that it is raining. Even if it is raining, one may lack the kind of causal contact with the rain that one needs in order to know that it is raining. The compositional structure of sentences and thoughts further constrains the ascription of knowledge, because the inferential processes in which subjects engage are sensitive to that structure: to interpret those processes as leading to knowledge, one must interpret them as valid inferences.

The more abundant ontology is, the more objects, properties and relations there are, the more scope there often is for an assignment of reference on which we know. Conversely, the sparser ontology is, the fewer objects, properties and relations there are, the greater the danger that we do not know on any assignment. But the correlation is imperfect, for a sparse ontology sometimes facilitates knowledge by reducing the number of wrong answers clustered around the right one and hard to distinguish from it. Knowledge maximization tilts the playing field in our favour without guaranteeing us victory.[19]

One might still fear that the knowledge maximization principle is overcharitable. Suppose, for example, that I can see only a small part of a ball, the rest of which is hidden by some obstacle. I judge of the ball 'It is red'. Un-

---

[18] An interpretation on which the agent believes that ticket *n* did not win might do better than one on which the agent believes that ticket *n* won, even though neither constitutes knowledge, if the former attributes more knowledge of chances to the agent than the latter does.

[19] Even if ontology is abundant, the nature of reference may also embody a bias towards more natural referents over less natural ones (see Lewis 1983a and, for the link with issues about intuition, Weatherson 2003). Might that counteract the bias towards knowledge? Not necessarily: the bias towards the natural in reference may reflect a bias towards the natural in knowledge.

known to me, the rest of the ball is green, so that the ball as a whole does not qualify as red. I falsely believe, and do not know, that the ball is red; what I know is that the visible part of the ball is red. Does knowledge maximization imply, falsely, that the visual demonstrative 'it' refers to just the presently visible part of the ball rather than to the whole ball? No. I am also disposed to judge 'It is a ball' and to infer from that judgement and 'It is red' to the conclusion 'There is a red ball'. If the visual demonstrative in 'It is a ball' is made to refer to the ball part, then the latter judgement is made false, and my knowledge of the visually presented ball that it is a ball is lost. If the visual demonstrative in 'It is a ball' is made to refer to the whole ball, then my inference to the conclusion 'There is a red ball' is fallacious. My use of the visual demonstrative is quite insensitive to the supposed shift in its reference. The two tokens of 'it' advert to a single mental file. But such confusion in the use of the file would undermine its capacity to function as a repository of knowledge, whether concerning the ball that it is a ball or concerning the ball part that it is red, and thereby subverts the point of imputing it. Thus the reference-shifting strategy fails to maximize knowledge. Of course, these considerations provide only a foretaste of all that is relevant. The next stage might be to question the reference of 'red' and 'ball'. But it is already appreciable that the holistic character of the considerations gives plenty of scope for the knowledge maximization principle to get the right answer, and arguably for the right reasons.

Another doubt about knowledge maximization concerns variants of the Celia/Elsie case above in which Emanuel knows independently that Elsie is F, G, H, ... However, he can still use 'she' as a visual demonstrative to refer to Celia in judging 'She is F', 'She is G', 'She is H', ..., thereby expressing false beliefs about Celia rather than knowledge about Elsie, because those judgements are not causally based on his independent knowledge of Elsie, and therefore fail to express that knowledge. Of course, in a further variant of the case, Emanuel makes the identity judgement 'She is Elsie', and then judges 'She is F', 'She is G', 'She is H', ..., on the basis of inference from the identity judgement and the premises 'Elsie is F', 'Elsie is G', 'Elsie is H', ..., so that his independent knowledge of Elsie is causally active in his reaching the conclusions. Even in that case, knowledge maximization still does not warrant assigning Elsie as the referent of the visual demonstrative 'she'. If knowledge is sensitive to differences in mode of presentation, and 'she' is associated with a visual mode of presentation, then the judgement 'She is Elsie' does not constitute knowledge; consequently, the further judgements derived from it also fail to constitute knowledge. On the other hand, if knowledge is not sensitive

to differences in mode of presentation, then assigning Elsie as the referent of 'she' merely makes the judgements 'She is F', 'She is G', 'She is H', ..., constitute the same knowledge as 'Elsie is F', 'Elsie is G', 'Elsie is H', ..., already constitute; no knowledge is gained. Moreover, that assignment also makes judgements such as 'She is standing in front of me' fail to constitute knowledge, whereas they do constitute knowledge on the assignment of Celia as the referent of 'she'. Hence the correct assignment (Celia) involves the ascription of more knowledge than the incorrect one (Elsie) does. Thus knowledge maximization does not involve the misinterpretation of such cases. Consideration of the inferential role of the perceptual demonstrative, as in the case of the ball, confirms that conclusion.

The knowledge maximization principle does not make the ascription of knowledge come too cheap. By contrast, Davidson's principle of charity would give good marks to an interpretation on which stone age people assented to many truths of quantum mechanics, if it happened to fit the compositional structure of their language. He tries to recover a plausible epistemology by extracting epistemological consequences from his principle of charity by appeal to the immunity from massive error that it is supposed to grant. That immunity is holistic: it is consistent with the falsity of almost any given one of our beliefs, given enough compensating truth elsewhere in the system. In particular, my belief that I have hands enjoys no immunity from error. The supposed immunity from massive error does not explain how I know that I have hands: likewise for most of what we ordinarily take ourselves to know. Davidson adds an appeal to causal constraints on reference in simple cases, but formulates the constraints too crudely to permit any straightforward connection with knowledge (Davidson 1991, 196-7). Even if my belief that P is caused by what it is about, I may fail to know that P because the causal chain is somehow deviant. When Davidson tries to explain how his principle of charity yields knowledge, he appears to rely on something like the pre-Gettier assumption that justified true belief is knowledge.[20]

A subtler attempt to extract knowledge from Davidson's principle of charity exploits beliefs that one knows. Very often, when one believes that P, one

---

[20] 'There is at least a presumption that we are right about the contents of our own minds; so in the cases where we are right, we have knowledge' (Davidson 1991, 194); 'Anyone who accepts perceptual externalism knows he cannot be systematically deceived about whether there are such things as cows, people, water, stars, and chewing gum. Knowing why this is the case, he must recognize situations in which he is justified in believing he is seeing water or a cow. In those cases where he is right, he knows he is seeing water or a cow' (Davidson 1991, 201). See also Davidson 1983.

also believes that one knows that P.[21] If one believes truly that one knows that P, then one does know that P. Does maximizing true belief therefore indirectly maximize knowledge too? The detour through second-order belief is unpromising. First, it depends on the assumption that the relevant agents are to be interpreted as believing that they know. Of course, we often believe that we know; for that matter, we often know. But the aim was to derive the conclusion that agents in general often know from a truth-maximizing principle of charity; that agents in general often believe that they know has not been derived from such a principle. Second, even granted that they believe that they know, Davidson's principle attributes no special status to beliefs of that form; an interpretation might sacrifice them all as false and still maximize true belief overall by making enough other beliefs true. Third, the account does not generate attributions of knowledge to simple creatures who lack the concept of knowledge and therefore cannot believe that they know; surely they can have knowledge without having the concept of knowledge.[22] Truth maximization lacks most of the epistemological rewards of knowledge maximization.

Quine endorses as a canon of translation the epistemological-sounding maxim 'Save the obvious' (Quine 1970, 82; compare Quine 1960, 59): do not interpret the natives as dissenting from obvious truths. On that basis he argues that apparent deviations in logic are mere artifacts of bad translation. Although this appears to invoke a knowledge-related standard of charity, like the principle of knowledge maximization, Quine insists on interpreting 'obvious' behaviouristically rather than epistemologically.[23] His intended maxim is that translation should preserve general assent. Without further argument, we can-

---

[21] The principle cannot be exceptionless, otherwise having any belief involves having infinitely many beliefs of increasing complexity.

[22] Davidson might be willing to deny that one can have knowledge without the concept of knowledge, for he denies that one can have beliefs without the concept of belief: 'Someone cannot have a belief unless he understands the possibility of being mistaken, and this requires grasping the contrast between truth and error – true belief and false belief' (Davidson 1975, 170). Whether or not he would extend it to knowledge, Davidson's argument is unconvincing, for it conflates *de re* and de *de dicto* readings. Grant for the sake of argument that, to believe that P, one must grasp the contrast between the state of affairs that P, which is in fact the condition for the belief to be true, and the state of affairs that it is not the case that P, which is in fact the condition for the belief to be false (the *de re* reading). Even so, Davidson does not explain why one must grasp it *as* the contrast between the condition for the belief to be true and the condition for it to be false (the *de dicto* reading), which is what he needs. Thus he leaves it obscure why a creature with the concept of negation could not have a belief without the concept of belief.

[23] 'I must stress that I am using the word "obvious" in an ordinary behavioral sense, with no epistemological overtones. When I call "1 + 1 = 2" obvious to a community I mean only that everyone, nearly enough, will unhesitatingly assent to it, for whatever reason; and when I call "It is raining" obvious in particular circumstances I mean that everyone will assent to it in those circumstances' (Quine 1970, 82).

not conclude that sentences that enjoy general assent are true, for we can assume neither that every sentence to which speakers of another language assent can be translated into English nor that every sentence to which speakers of English assent is true – naturally, it is hard for us, as speakers of English, to produce a counterexample. Like Grandy's principle of humanity, Quine's maxim on its behavioural reading tends to project our design faults onto others. For example, it discourages us from translating a sentence to which the natives universally assent by a simple logical truth from which many speakers of English dissent through intellectual confusion. On an epistemological reading, the maxim is not vulnerable to that criticism, for confused speakers can dissent from what is obvious.

Knowing is itself an intentional state. How much a given subject would know on a given assignment of reference is an intentional question. That would be a serious problem if we were trying to use the principle of knowledge maximization in a reduction of the intentional to the non-intentional. But the present aim is different. The issue was whether there is any theoretical basis to predict a general, highly defeasible tendency for beliefs to be true. We have found some reason to suspect that the nature of reference is to work in a knowledge-maximizing way. If so, the nature of reference grounds a mild tendency for beliefs to constitute knowledge. Since knowledge entails truth, the nature of reference grounds a mild tendency for beliefs to be true, even though maximizing knowledge is not equivalent to maximizing true belief.

On a more internalist proposal, the nature of reference is to maximize justified belief rather than knowledge, where justified beliefs can be false. But such internalism makes the bearing of reference on justification obscure. Suppose that I have a few factual memories of a brief acquaintance, which I express using the pronoun 'he'. The assignment of one reference rather than another to 'he' seems to make no difference to the internalist justification of my memory beliefs; it makes an obvious difference to whether they constitute knowledge. Similarly, internalist considerations of justified belief are much less likely than externalist considerations of knowledge to explain why the silly interpretation Int* in the previous section is worse than the sensible connection Int, for the permutation of contents preserves internal connections but not knowledge. The external involvement of knowledge exactly suits it to constrain reference.

Is it surprising for reference to maximize knowledge? Reference concerns what mental states and acts are about. Knowledge is one mental state amongst many. Why should it play a privileged role in determining what all of them are about? One answer is that knowledge is not just one mental state amongst many. A creature that is not aware of anything at all has no mental life. It lacks

genuine intelligence. Although intelligent life does not consist solely of aware-ness, it is intelligent only because appropriately related to awareness of some-thing. But to be aware is to know: one is aware that P if and only if one knows that P. Intelligent life is life appropriately related to intelligent action, and in-telligent action is action appropriately related to knowledge. In a paradigm of intelligent action, given a desire that P, one knowingly does A, knowing that if one does A then P. One can believe that one does A and that if one does A then P, even truly, without knowing, but in such cases the action is defective; they are to be understood in relation to non-defective cases. The function of intelligent action involves the application of knowledge to realize the agent's ends. In unfavourable circumstances, only mere beliefs are available, and in-tentional action does not function properly, although with good luck it may still achieve the desired end, just as other defective processes sometimes issue in the intended product.[24]

When conditions are unfavourable, the agent is not in a position to know anything much, just as a victim of total paralysis may not be in a position to do anything much. Intentional action may be limited to pursuing a line of thought. For a brain in a vat, both knowledge and action may shrink to the in-ternal: but that pathological case does not reveal their underlying nature, for it does not show them to be equally shrunken in more normal cases. Rather, the pathological cases are parasitic on the normal ones.

Given the central role of knowledge in intelligent life, the intimate relation between knowledge and reference is hardly surprising. Reference maximizes knowledge because it serves knowledge and should therefore impose no in-dependent limitation on it. Consequently, the nature of reference is not to be strictly neutral between true and false belief, nor even between knowledge and mere true belief.

---

[24] See Williamson 2000 for more on the associated conception of mind and knowledge. The idea that all thinking qualifies as such by being appropriately related to knowing was advo-cated by John Cook Wilson (1926, vol. I, 35-40, also for the view that knowledge is indefin-able). He defends a neo-Aristotelian version of common sense realism on which ordinary lan-guage has a central role in metaphysics. Of the 'examination of the meaning of grammatical forms' and the consideration of 'certain distinctions of the kind called metaphysical' he says 'The two investigations are necessarily connected with one another; for since the sentence or statement describes the nature of objects and not any attitude of ours to the objects described, in the way of apprehension or opinion, its meaning is wholly objective, in the sense that we have already given to objective. That is, it is about something apprehended, in the case of knowledge for instance, and not about our apprehension of it' (1926, vol. I, 149). In some very general respects the present paper belongs to a tradition that runs from Cook Wilson to Prichard and others, then to Austin and later to John McDowell; see Marion 2000. It hardly needs saying that there are also very significant differences between these philosophers.

6. Knowledge and intuition.

An externalist picture of the mind has been sketched, with the broadest strokes,
on which the nature of reference embodies a bias towards the evaluation of
judgements as knowledgeable, not merely true. Given our tendency to judge
as we are inclined to judge, the bias extends to conscious and unconscious in-
clinations to judgement. That all helps put the burden of proof on sceptics
about judgement to argue that their radical scenarios deserve to be taken more
seriously than do the radical scenarios for scepticism about perception. Al-
though we can allow that scenarios of both sorts are metaphysically possible,
much more than that is needed to justify serious doubt. The burden of proof
on the sceptic about judgement is particularly heavy when the proposed sce-
narios make vast ranges of common judgements false or otherwise unknowl-
edgeable, as many of them do.[25]

Someone might respond thus:

Granted, when we are consciously inclined to judge that P, we often
       but not always know that P. That we are consciously inclined to judge
       that P should therefore be treated as good but defeasible evidence for
       the claim that P. It is just one more part of the total body of evidence
       on which philosophical theories should be evaluated.

What # perversely ignores is the evidential role of the fact that P itself, as op-
posed to that of the fact that we are consciously inclined to judge that P. After
all, if we do know that P, would it not be negligent not to use that knowledge
in evaluating a philosophical theory to which it is relevant? Philosophy is hard
enough already: why should we make it even more difficult by forbidding our-
selves to bring some of our knowledge to bear? We are not obliged to fight
with one arm tied behind our back.[26]

Someone might reply on behalf of # that, if we know that P without know-
ing that we know that P, then the knowledge does not really help. But that re-
sponse is doubly inadequate. First, it gives no more reason to deny that we
know that we know that P than to deny that we know that P in the relevant
cases. Although we cannot expect to have infinitely many iterations of knowl-
edge, for more than computational reasons (Williamson 2000, 114-34), that
general point merely shows that we must sometimes simply apply our knowl-

---

[25] The case of folk physics, mentioned in section 1, does not constitute a sceptical sce-
nario, for folk physics plays a role in generating much knowledge of particular facts about our
environment.
    [26] See Williamson 2000, 184-208 for defence and development of the conception of our
total evidence as everything that we know.

edge, without first checking whether we know, for otherwise we get stuck in an infinite regress of checks. That is the second problem for the envisaged defence of #. It gave us no reason to think that we are entitled to rely on the premise that P in philosophical discussion only if we know that we know that P.

When we know, there is presumably something non-trivial to be said about how we know. But we may know that P, and even know that we know that P, without knowing how we know that P. Perhaps we know that we know the truth of some logical or mathematical axioms without knowing how we know their truth. Similarly, the epistemic load that considerations of theoretical elegance and simplicity bear in theoretical physics seems as indispensable as it is hard to explain. But for many judgements that philosophers are tempted to classify as intuitive, the question 'How do you know?' is not especially puzzling. There is no special mystery as to how we know that there are mountains in Switzerland. We can say how we know, typically by describing the process by which we acquired the knowledge, without having to convince the sceptic who doubts that we know.

The knowledge maximization principle is not itself intended as an answer to the question 'How do you know?'. The knowledge maximized may have been acquired by quite familiar means of perception, memory, testimony and inference, and perhaps whatever means we use to learn about modal facts. The proper response to scepticism about judgement is not to postulate intuition as a separate means to knowledge to underpin all the others but rather to challenge the sceptical idea that they need such underpinning. Its supposed function would be to rule out the scenarios that motivate scepticism about judgement. But a good answer to the question 'How do you know that P?' need not specifically address far-fetched sceptical scenarios for the belief that P, since specific consideration of them need not have occurred when the knowledge was acquired. Knowledge maximization is a factor, typically unnoticed by sceptics about judgement, that makes their scenarios more far-fetched than they realize.

The sceptic may go on to ask 'How do you know that you are not in this sceptical scenario for the belief that P?'. That P entails that we are not in the sceptical scenario, so we might answer 'By knowing that P and deducing that we are not in the sceptical scenario'. Many philosophers feel uneasy about that answer; they fear that it somehow begs the question against the sceptic, especially when the answer to the question 'How do you know that P?' ignored the sceptical scenario. Whatever the reason for that unease, it is in no way specific to scepticism about judgement. It arises equally for scepticism about perception. We do not let the sceptic about perception's scenarios panic us into serious scepticism about perception. We should be equally reluctant to let the scep-

tic about judgement's scenarios panic us into serious scepticism about judgement. That the sceptic about perception wears the traditional garb of the philosopher while the sceptic about judgement dresses up in a scientist's white coat should not blind us to the underlying structural similarity of their arguments. Much remains to be done in the diagnosis and treatment of those arguments, but we are not obliged to be sceptics until the work is completed.

In less counterfactual mode, the sceptic tries to induce a crisis of confidence in present common sense by pointing towards a present or future scientific outlook that stands to present common sense as the latter stands to a stone age outlook. But the analogy rebounds against scepticism about judgement. For although it is plausible that stone age people had many false beliefs about the general nature of the world, it is equally plausible that they had significant knowledge of their local environment. We have no good reason to think that there was ever a stage at which humans had beliefs without knowledge. Knowledge maximization plausibly suggests that our ancestors had some primitive knowledge as soon as they had some primitive beliefs; it is not as though archaeology suggests otherwise. Similarly, if it is plausible to attribute primitive beliefs to some non-human animals, it is equally plausible to attribute to them some primitive knowledge of their local environment.

What are we to make of empirical evidence for clashes of intuition? There is empirical evidence that eye-witness reports of events are alarmingly unreliable. Such evidence gives us good reason to be somewhat more cautious than we might otherwise have been in our use of perception, memory and testimony; it does not give us good reason to discount all evidence reached through perception, memory or testimony (including the empirical evidence about eye-witness reports). Similarly, clashes of what philosophers call 'intuition' give us good reason to be somewhat more cautious than we might otherwise have been in our applications of philosophically significant concepts; they do not give us good reason to discount all evidence reached through such applications. It would be rash to assume without evidence that skills at eye-witness reporting are evenly distributed across individuals or social groups; it would be equally rash to assume without evidence that skills at applying abstract concepts to complex examples are evenly distributed across individuals or social groups. Legal or philosophical training may even improve such skills: although it is harder to find an uncontroversial test of that hypothesis than of one about the effect of training in eye-witness reporting, that is not evidence against the hypothesis.

Consider this analogue of # concerning evidence for or against empirical theories:

Granted, when we are inclined to judge perceptually that P, we often but not always know that P. That we are inclined to judge perceptually that P should therefore be treated as good but defeasible evidence for the claim that P. It is just one more part of the total body of evidence on which empirical theories should be evaluated.

What ## perversely ignores is the evidential role of the fact that P itself, as opposed to that of the fact that we are inclined to judge perceptually that P. After all, if we do know that P, would it not be negligent not to use that knowledge in evaluating an empirical theory to which it is relevant? It would not advance science to insist that scientists' evidence cannot include the fact that 19 out of 20 rats fed the substance died within 24 hours, but only the fact that the scientist was perceptually inclined to judge that 19 out of 20 rats fed the substance died (only the former fact leads itself to statistical analysis). Such claims about inclinations to judgement are not infallible. Indeed, they are less amenable to public checking by the scientific community than are claims about what the outcome of the experiment actually was. Of course, it may later turn out that a disgruntled lab technician fed the rats the wrong substance, but the proper response to such remote possibilities is to backtrack if one of them is found to obtain, not to make a futile attempt in advance to identify evidence for which backtracking will not be required in even the remotest eventualities.

In philosophy as in empirical science, our evidence does not consist of facts with respect to which we are infallible. Ordinary knowledge is enough. We have no general guarantee against the possibility that we did not know something that we thought we knew. In philosophy, the evidence is even more contested than in empirical science. As we saw in section 2, even psychologized evidence is philosophically contentious. The philosopher's predicament is comparable to that which would face the empirical scientist if accusations of falsified evidence were vastly more common in science than they actually are. Whatever the discipline, when someone disputes the evidence, it is often better to look for common ground on which to pursue the argument than to ride roughshod over the objections. For that temporary purpose, we may refrain from treating the disputed evidence as evidence. But that does not entail that it should never have been treated as evidence in the first place. Moreover, the search for common ground can be taken too far, especially with an irresponsible opponent who does not scruple to challenge any inconvenient evidence. A sufficiently ruthless sceptic can challenge everything that we offer as evidence, by always demanding a proof; that should not drive us to suspend all our evidence. At some point we are entitled to hold on to what we know, and apply it.

In philosophy, our evidence consists of a miscellaneous mass of knowledge, expressed in terms of all kinds, some from ordinary language, some from the theoretical vocabulary of various disciplines. Some of it consists of knowledge about our own mental states; most of it does not. Anything we know is legitimate evidence. Inevitably, philosophers make mistakes, sometimes treating as known what is unknown, or as unknown what is known. Since philosophy is unusually tolerant of scepticism, the extent of our evidence is unusually contested in philosophy. Our fallibility about our evidence is unavoidable; we have no alternative but to muddle through as best we can.

Metaphilosophical talk of intuitions obscures our real methodological situation in at least two ways. First, it feeds the methodological illusion of an incontestible starting-point, if not of intuited facts, then of facts as to what we intuit. There is no such starting-point; evidence can always be contested. Second, it conceals the continuity between philosophical thinking and the rest of our thinking. So-called intuitions involve the very same cognitive capacities that we use elsewhere, but deployed in contexts in which scepticism about judgement is salient. If we want to identify what is genuinely distinctive in philosophical thinking, we should stop talking about intuition.[27]

REFERENCES

BEALER, G. 1998: 'Intuition and the autonomy of philosophy', in DePaul and Ramsey 1998.
BEALER, G. 2002: 'Modal epistemology and the rationalist renaissance', in Gendler and Hawthorne 2002.
DAVIDSON, D. 1974: 'On the very idea of a conceptual scheme', *Proceedings and Addresses of the American Philosophical Association* 47, 5-20. Reprinted in Davidson 1984, to which page numbers refer.
DAVIDSON, D. 1975: 'Thought and talk', in S. Guttenplan, ed., *Mind and Language*, Oxford: Oxford University Press. Reprinted in Davidson 1984, to which page numbers refer.
DAVIDSON, D. 1977: 'The method of truth in metaphysics', in P. A. French, T. E. Uehling, Jr., and H. K. Wettstein, eds., *Midwest Studies in Philosophy, 2: Studies in the Philosophy of Language,* Morris: The University of Minnesota. Reprinted in Davidson 1984, to which page numbers refer.

DAVIDSON, D. 1983: 'A coherence theory of truth and knowledge', in D. Henrich, ed., *Kant oder Hegel?*, Stuttgart: Klett-Cotta. Reprinted with 'Afterthoughts' in Davidson 2001.

DAVIDSON, D. 1984: *Inquiries into Truth and Interpretation.* Oxford: Clarendon.

DAVIDSON, D. 1991: 'Epistemology externalized', *Dialectica* 45, 191-202. Reprinted in Davidson 2001, to which page numbers refer.

DAVIDSON, D. 2001: S*ubjective, Intersubjective, Objective.* Oxford: Clarendon Press.

DEPAUL, M., and RAMSEY, W. (eds.) 1998: *Rethinking Intuition: The Psychology of Intuition and its Role in Philosophical Inquiry,* Lanham, Maryland: Rowman and Littlefield.

DUMMETT, M. A. E. 1975: 'The philosophical basis of intuitionistic logic', in H. E. Rose and J. C. Shepherdson, eds., *Logic Colloquium* '73, Amsterdam: North-Holland.

FEYERABEND, P. 1978: *Science in a Free Society*. London: NLB.

GENDLER, T. SZABÓ, and HAWTHORNE, J. (eds.) 2002: *Conceivability and Possibility,* Oxford: Clarendon Press.

GETTIER, E. 1963: 'Is justified true belief knowledge?', *Analysis* 23, 121-3.

GOLDMAN, A., and PUST, J. 1998: 'Philosophical theory and intuitional evidence', in DePaul and Ramsey 1998.

GOODMAN, N. 1965: *Fact, Fiction and Forecast.* New York: Bobbs-Merrill.

GRANDY, R. 1973: 'Reference, meaning, and belief', *Journal of Philosophy* 70, 439-52.

HARMAN, G. 1986: *Change in View: Principles of Reasoning,* Cambridge, Mass.: MIT Press.

HORGAN, T. 1995: 'Transvaluationism: a Dionysian approach to vagueness', *Southern Journal of Philosophy* 33, supplement, 97-126.

KLEIN, P. 1986: 'Radical interpretation and global skepticism', in LePore 1986.

LEPORE, E. (ed.) 1986: *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, Oxford: Blackwell.

LEWIS, D. 1974: 'Radical interpretation', *Synthese* 23, 331-44. Reprinted with 'Postscripts' in Lewis 1983a, to which pages number refer.

LEWIS, D. 1983a: *Philosophical Papers*, vol. 1, Oxford: Oxford University Press.

LEWIS, D. 1983b: 'New work for a theory of universals', *Australasian Journal of Philosophy* 61, 343-77.

MARION, M. 2000: 'Oxford realism: knowledge and perception', parts I and II, *British Journal* for the History of Philosophy 8, 299-338 and 485-519.

MCCULLOCH, G. 2003: *The Life of the Mind: An Essay on Phenomenological Externalism,* London: Routledge.

MCGINN, C. 1986: 'Radical interpretation and epistemology', in Lepore 1986.

MOORE, G. E. 1925: 'A defence of common sense', in J. H. Muirhead, ed., *Contemporary British Philosophy* (second series), London: George Allen & Unwin.

NAGEL, T. 1986: *The View from Nowhere,* Oxford: Oxford University Press.

Nozick, R. 1981: *Philosophical Explanations.* Oxford: Oxford University Press.

PUST, J. 2001: 'Against explanationist skepticism regarding philosophical intuitions', *Philosophical Studies* 106, 227-58.

QUINE, W. V. 1951: 'Two dogmas of empiricism', *Philosophical Review* 60, 20-43.

QUINE, W. V. 1960: *Word and Object*, Cambridge, Mass.: MIT Press.

QUINE, W. V. 1970: *Philosophy of Logic,* Englewood Cliffs, NJ: Prentice-Hall.

RAWLS, J. 1971: *A Theory of Justice,* Cambridge, Mass.: Harvard University Press.

VAHID, H. 2003: 'Varieties of epistemic conservativism', *Synthese,* forthcoming.

VAN INWAGEN, P. 1995: *Material Beings*, Ithaca, NY: Cornell University Press.

VAN INWAGEN, P. 1997: 'Materialism and the psychological-continuity account of personal identity', in J. Tomberlin, ed., *Philosophical Perspectives, 11: Mind, Causation and World,* Oxford: Blackwell.

WEATHERSON, B. 2003: 'What good are counterexamples?', *Philosophical Studies* 115, 1-31.

WEINBERG, J., STICH, S., and NICHOLS, S. 2001: 'Normativity and epistemic intuitions', *Philosophical Topics* 29, 429-460.

WILLIAMSON, T. 2000: *Knowledge and its Limits,* Oxford: Oxford University Press.

WILLIAMSON, T. 2003: 'Understanding and Inference', *Aristotelian Society* sup. vol. 77, 249-293.

WILSON, J. C. 1926: *Statement and Inference,* 2 vols., Oxford: Clarendon Press.