Saving belief from (internalist) epistemology

Adam Morton

Why do herbivorous animals gather in herds or schools? According to biologists, the principal reason is because a thousand eyes are better than two. By pooling information about their surroundings they can increase the probability of detecting the presence of predators and other dangers. Humans too share the information they gather. We do it by speech. Three things are thus intimately related in human life: the acquisition of beliefs on the basis of evidence – one member of the band sees a wolf – their expression in language – that person calls out to others – and the acquisition of beliefs about the environment and about others – others think "there's a wolf", "Hugo thinks there's a wolf", "Hugo is trying to get us to think there is a wolf."

The presence of language makes the human situation much more complex than that of other animals in this respect. Other animals do not assess the warning signal of a herd-mate and decide whether it was mistaken or deceptive. But we do. We have to assess the *justification* and *reliability* of one another's *beliefs* in order to use one another as sources of information. The three italicized terms are crucial here, but they are also subtle and treacherous. What, in particular, is it to believe something? Philosophers expend a lot of effort on trying to define belief, and psychologists expend a lot of effort trying to understand how a child comes to have a concept of belief. Both philosophers and psychologists usually assume that to understand the concept of belief one has to understand when a belief is correctly acquired: these are the ideal cases in terms of which we can explain irrational or otherwise mis-formed belief. If this assumption is correct, epistemology is basic to belief-ascription: to understand belief we must understand justification and reliability.

I shall not question the link between epistemology and belief. The point of this paper is more limited: ideas about the acquisition of belief that have lost credibility in epistemology survive in the philosophy of mind and in developmental psychology. Once we realize how many options we have when understanding the reasons for which beliefs may be held, the scales fall from our eyes and we see the obvious, that conceptions of what it is to believe something and what it is to ascribe a belief to someone current in some parts of philosophy and psychology are extremely implausible. We only believed them because we thought there was no alternative.

Which epistemological positions and which features of accounts of belief are

we talking about? The exact claims will emerge more cautiously in the course of the paper, but the first crude statement is that once we see the problems with the old *internalist* orthodoxy in epistemology then we begin to see problems with the dominant *functionalist* understanding of belief, and in particular with ideas about belief attribution that accompany it. And if we embrace the alternative *externalist* epistemologies then we begin to see an outline of a non-functionalist account of belief and belief attribution. All these labels will need explanation; I begin with internalism and externalism. Before doing that, though, it might help to say that this paper contains both negative and positive claims. The central negative claim is one that I think is very hard not to accept: when we abandon internalism functionalism looks very dubious. The central positive claim is that externalism can found an alternative to functionalism. That is a much more conjectural claim, explored in the last sections of the paper. I present it as enlightening conjecture rather than as logical deduction.

Internalism/externalism: some basic background The topic is how beliefs *should* be acquired: how we would get our beliefs if we were not subject to mistakes, bad reasoning, illusions, irrationality and the like. Epistemologists focus on rational beliefs acquired under good conditions because they want to criticize existing bodies of belief or articulate ideals for forming less imperfect bodies of belief. But the normative can often be taken as an outline of the normal: we take belief to be to a first approximation to what would ideally be believed in the believer's circumstances. (Bear in mind a distinction, though: the approximation is more plausible as a description of what we *take* belief to be, in ordinary life, than what it *is*.) But humans are so un-ideal: that list "mistakes, bad reasoning, illusions, irrationality" is so varied that the subsequent "and the like" ought to make us stop and think. Is there a common theme to the many ways our attempts at forming true beliefs, about things that we need to understand, can go wrong?

According to the internalist tradition in epistemology the relevant theme is good and bad reasoning. We are capable of controlling our reasoning and bringing it nearer and nearer to the ideal. (As Descartes put it, once he became aware of his core cognitive limitation, that of limited capacity for continuous concentration, all that was needed was "que je ne manque jamais de m'en resouvenir, toutes les fois que j'en aurai besoin, et acquérir de cette façon l'habitude de ne point faillir."[1] Easy!) So the effective focus for epistemology becomes that of providing descriptions of good transitions between evidence and belief for us to follow, and the practical focus in the everyday understanding of belief is what the evidence for a belief is and what kinds of reasoning can be used to get it. We understand what someone thinks by seeing what evidence is available to her and what inferences

---

[1] *Meditation 4*

she made from this evidence. If you see your neighbor approach the spot where her car should be but is not, and then rush back into her house, you suppose that she has noticed its absence, and reasoned to the conclusion that it has been stolen, and gone to call the police.

Speaking of material for inference as "evidence" suggests that beliefs are inferred from non-beliefs, or from special intrinsically initial beliefs. That is the traditional internalist assumption, common to both Cartesian and empiricist versions. The result is what is usually called foundationalism. Internalism can use quite different assumptions, though. Instead of taking beliefs to be inferred from evidence we can take beliefs to be inferred from other beliefs. More specifically, a belief can be taken to be inferred from the person's total body of previously held beliefs, which will include the beliefs that would intuitively be described as stating the relevant evidence available to the person. This is the holist or coherentist version of internalism. In its most convincing version, due to Gilbert Harman, a *change* in a person's total body of beliefs results from reasoning from her total body of previously held beliefs. (Fuller versions of these standard positions, with more details and more qualifications, can be got from the standard textbooks.[2])

Flawed reasoning is not the only thing that produces false beliefs. Defective inputs, for example perceptual illusions, can doom the most careful inference. But a person is not to blame for the beliefs she derives from evidence she had no reason to believe was misleading. The blame here is the stigma of irrationality: beliefs based on illusions or other misleading evidence are often rational. There are two internalist assumptions here. The first is that the central evaluative concepts we apply to beliefs are a matter of evaluating their rational connections to evidence or to other beliefs, where this evidence or belief is also represented in the mind of the person in question. It is rationality that matters, and rationality is a matter of connections between states of a single person. The second assumption is that the states in question are potentially knowable by the person in question, and the transitions between them to some extent under their control. Imputations of irrationality can thus have a corrective effect on our belief-acquisition habits. Any internalist epistemology will subscribe to some form of both these assumptions. (A modern internalism will obviously want to hedge or dilute the second one to some extent.)

Neither of these assumptions is obviously true. Many would consider them both obviously false. The question is hard to settle in that it is in part a matter of emphasis and value: the first assumption says that our 'central' evaluations – the most important ones – apply to internal states of an individual, and the knowability

---

[2] A sophisticated, though now slightly out of date, textbook is Dancy 1985. More recent material is covered in Greco and Sosa 1999. For a survey from the point of view of the philosophy of science see Morton 2000. For an uncomplicated elementary exposition see Morton 2002. For Harman's version of holism see Harman 1986.

postulated in the second assumption applies only to the states which are the objects of these central evaluations. But to see how one can profitably differ from internalism consider the variety of beliefs that are based on processes that we cannot even try to tune, correct, or control in the deliberate ways that we manipulate inference from evidence to theory. The best example is our beliefs about the moods and characters of people we meet face to face. To a large extent these beliefs are like basic perceptual beliefs – we just know that the person we are talking to is annoyed, or that they are trying to tell us something important, or not taking us seriously, or whatever. We have no real insight into how we get these beliefs, though they are obviously the result of complex mental processes. And we rely on them for practically all of our activities. An internalist might thus classify them as basic beliefs, evidence for other beliefs but not themselves to be evaluated. And indeed if I mistakenly think a person is annoyed at me when in fact she is upset about something I know nothing about, my belief is not irrational, though it may be a sign that I am not very perceptive about others. But this is an embarrassing admission: our lives depend on our being perceptive about one another, and evaluation of our capacities in such respects is hardly a peripheral or unimportant matter. An epistemology that focused on questions of how reliable our belief-forming processes are, in various situations, is obviously asking questions to which we need systematic and well-supported answers. But the processes in question are most often not ones to whose workings we have any conscious access. We cannot tell "from the inside" whether our beliefs are formed by reliable processes. But we care a lot whether they are, and we care a lot whether others are reliable sources of information on various matters.

One we focus on the reliability of belief-forming processes rather than the correctness of reasoning, we are taking the externalist perspective. It consists in asking third-person normative questions about belief, variations on "when can we trust what person p thinks about topic t?" One variation "when can I trust my own beliefs about t?" includes the internalist question "when is my reasoning correct?" as a special case. The obvious question that then arises is how much of the normative discourse that can be framed in externalist, third-person, terms can be translated back into internalist, first-person, terms. The consensus among contemporary epistemologists is that the answer is "not much" and that this accounts for some of the frustrations of the subject's history.

From internalism to functionalism We will return to externalism below. For now suppose that we are internalists, and see what conception of belief we will be pushed towards. From an internalist point of view beliefs are evaluated, as justified or unjustified, in terms of the quality of the inferences on which they are based. There is little point doing this if substantial numbers of beliefs are not so based, so

we assume that we can associate with most beliefs a trace of inferences and intermediary beliefs which shows the belief's provenance. Suppose now that we are foundationalist internalists. (I will turn later to what happens if we are internalists of the holistic kind.) Then we assume that the trace leads back to primary evidence about the world, either to something perceptible by all or private to the individual.

This is the core idea that I believe implicitly underlies many accounts of belief, in both psychology and philosophy. In effect, though the idea is rarely stated in these terms, it makes inference primitive and defines belief as what a person infers from what she perceives, is told, or in some other way acquires as basic information. The idea emerges when, for example, we consider a child who sees that the evidence available to a person leads by a simple inference to a false belief, but does not ascribe that false belief to the person. We find this surprising and paradoxical, a sign that the child does not really have the concept of belief. For after all belief is what you get by inference from what you perceive. (The most disturbing fact, from this foundationalist point of view, is not that the child attributes the wrong belief, but that she attributes any belief at all. For one would expect either that she would understand the evidence-based reasoning involved, and attribute the right belief, or that she would not grasp it and thus fail to attribute any belief at all.) And it makes intelligible our feats of attribution. Consider what is going on when one person attributes to another the belief that most cancers are not caused by viruses. How on earth does she do it? She is relating the other to objects and properties that cannot be perceived and connecting these remote entities on the other's behalf in complex ways, without access to the words that run through the other's consciousness. The answer that foundationalism suggests is that the attributer must be aware of the sources of the other's belief - the information available - and must be aware of the patterns of reasoning that the other has mastered, and then follows these lines of reasoning to the target conclusion. (One can follow from above, describing the reasoning and thinking about it, or from ground level, reproducing the reasoning to simulate the belief-attributing process. For these purposes it does not matter which.) In this way the mystery might seem to dissolve. We can attribute complex beliefs about remote entities because we can understand their epistemic grounds.

This is a broadly functionalist conclusion: a belief is to be understood in terms of its inferential connections with other states, particularly other beliefs. Add the assumption that this is the only grasp we have of belief, and what we get just is functionalism: beliefs are whatever plays the role of being the things that inference connects. Assuming functionalism, we can understand not only how people manage the attributional feat, but also what it is that the feat is attributing. We can say what someone means when they say "James believes that most cancers are not caused by viruses." What it means is that James is in a state which tends to be

related to other states in specifiable regular ways, prominent among which are the tendency to be produced by reasoning like that which connects it to its evidential basis, and the tendency to produce other similar states for which it itself is evidence.

Epistemic foundationalism, attribution-by-evidence, and functionalist accounts of the nature of belief (and by implication other states) are mutually supportive. The support is not a matter of logical necessity: one could without contradiction hold and deny any combination of the three. Yet if we accept all three we have a combination that makes sense together: a few additional assumptions turn foundationalism into attribution-by-evidence, and the resulting combination fills out the details of a functionalist schema in an intelligible way.

For all that, the crucial element in the combination, epistemic foundationalism, is a bad idea. The familiar criticisms of it focus on its failings as an account of human knowledge. It puts implausible and perhaps impossible constraints on the nature of the perceptual basis; it creates dilemmas about inductive reasoning; it rules out most patterns of actual scientific and common sense reasoning. I will not spell these out, as they are familiar in the epistemological literature of the past forty years. (See the material cited in footnote 2.) But these failings are closely related to its failures as an account of belief attribution. The short way to describe these failings is to say that if foundationalism is to ground belief, attribution must be taken deterministically. That is, given a person and an evidential situation there must be a single inevitable trace of beliefs that the person will derive from the situation. If not, knowing the situation does not tell us what the person believes.

But this is obviously not the case. It is arguable that an absolutely perfectly rational agent would gather a complete inventory of available evidence and would then derive *the* conclusions that are justified by it. (Arguable, but very far from clearly true.) But for any creature even remotely like a human being the beliefs that are formed in any situation depend delicately on many variables. There are the aspects - always incomplete - of the available evidence that are absorbed, the order in which individual items of evidence are absorbed, the structure of the prior beliefs that affect the interpretation of the evidence, the particular prior beliefs whose relevance is realized, the explanatory hypotheses that are formed, and other more idiosyncratic factors (whether the person is paying attention, whether they are intellectually energetic, their tendencies to self-deception, and so on indefinitely). It is true that in most situations there are beliefs that most normal human beings would derive from the situation (that the sun is shining, that the person on the sidewalk is dead). But that fact should not hide the more significant one that in all of these situations the majority of the beliefs that are formed will not be predictable. Knowing what evidence is available to someone does not tell you much about what they believe.

One response to these problems might be to go holist. Though beliefs are not even nearly determined by evidence alone, they are rather less undetermined by evidence plus prior belief. Given the whole complex of a person's beliefs, the ways that person can reasonably interpret the available evidence, and the conclusions they can reasonably then draw from it, are considerably more constrained. Now, while this is indeed a step towards a better epistemology (or so I would argue, the issue is definitely not beyond controversy) it does not help us understand attribution. There are two reasons. The most obvious reason is that the process would be circular. In order to attribute a belief one has to have attributed many other beliefs. So given a person and a situation one would have no idea what beliefs to attribute. Attribution via epistemic holism could at most be an adjunct to more robust methods, filling in a few additional beliefs once the main body of a person's beliefs were known. Even this would be a pretty dubious business, though, since a new body of beliefs is rationally determined, according to epistemological holism, by the *total* body of prior beliefs plus evidence. So in order to get an adequate grasp of the changes it is reasonable for a person to make in her beliefs, given the evidence that has become available, one would have to know a far greater proportion of her beliefs than is normally possible[3].

At this point the second, somewhat less obvious, problem arises. The problem is that we have almost no effective grasp of the rational transitions from one total body of belief to another. We know what the desiderata of such transitions are. They should maximize explanatory and logical coherence and reduce tensions, inconsistencies, and anomalies. But that tells you as little about what changes another person will make in their beliefs as it tells you about what changes you should make yourself. (There is a deep logical fact here, which applies to just about any form of cognition. Testing a set of sentences for consistency is a much harder task than extracting consequences one by one from a finite set of premises. Adherence to a set of epistemic principles expressed in terms of consistency or related concepts will not be naturally described as following a rule.) So though epistemic holism may be the truth about how we should change our beliefs, it gives us very little hold on the beliefs that any particular person may have acquired.

A quick point about functionalism Grant that when we reject epistemic internalism we lose a motive for accepting functionalism as an account of the nature of belief. Does it follow that functionalism is *wrong*? Well, at any rate

---

[3] Donald Davidson's account of belief - as revealed in for example Davidson 1984 - take holistic constraints to determine the beliefs that can be attributed to a person. But whether or not this position is correct – and the majority of philosophers have grave doubts about it – it is not intended as an account of how belief attribution could actually ever proceed, but as a constraint on the attributions that could be true of a person.

functionalism becomes very much harder to state. It becomes rather unclear what it is that we might be accepting. For according to functionalism beliefs are mental states that bear the right kinds of relation to other states, to perceptual input, and to behavioral output. What are the right relations? If foundationalism were true there would be definite relations to perceptual input, which could potentially be stated and which even unstated could be part of our grasp of belief. Once we abandon these we retreat to holism, but then the relations between bodies of beliefs become crucial. And these are utterly mysterious. We are left with something like "beliefs are those states which evolve in the direction of maximum explanatory and logical coherence." That is not really an adequate characterization, for it does not tell us when a particular state is a belief. It is as if we characterized a moose as any kind M such that members of M interact with other members of M to produce more members of M. The equation is not false of moose, but it is true of many other kinds too. There are far too many unintended solutions. We can be pretty sure that the same will hold with the analogous move for belief. And indeed conjectures, hypotheses, and expectations are objects of inference just as beliefs are. It has been argued, notably by Richard Foley, that some of these other states are better characterized in terms of their inferential relations than belief is[4]. So the dilemma facing the functionalist is this: either we incorporate a foundationalist epistemology in the functional characterization of belief, in which case the account is false, or a holist epistemology, in which case the account is vacuous.

The externalist alternative   So abandon epistemic internalism, as a way of grasping belief. What are we left with? In contemporary epistemology internalist considerations are complemented with externalist ones. (The two do not inevitably exclude one another, inasmuch as rationality and reliability are both important. On some topics, though, notably the definition of the concepts of justification and of knowledge, internalist and externalist rivals are often presented as alternatives.) Externalist considerations focus not on what reasoning would be rational for a person but on what beliefs and other states she would form if she were representing her environment accurately and reliably[5]. Visual perception provides a good example. When people are suitably placed with respect to suitable solid objects they normally see them. In so doing they normally form beliefs about the objects' locations and salient properties. They do not do this by explicit rationally evaluable reasoning. In fact, they normally do it even when there is strong evidence against the resulting belief. (A reasonable, but false, conviction that you

---

[4]   See Foley 1993.
[5]   I am inevitably simplifying, and glossing over important problems. For more careful and worried presentations see Goldman 1999, and chapter 4 of Williamson 2000.

are hallucinating does not stop you seeing.) We rely on what other people perceive because their perception is normally reliable, and we temper our reliance with estimates of their reliability. So if a very near-sighted person tells you that the bus half a mile away is a number 67B you don't leave the bus shelter until you can read the number for yourself. We also rely on reports of other people's perceptions and their other environment-tracking processes. So when someone says to you "Susanna sees the bus turning the corner" or "Susanna thinks the bus is turning the corner" then, provided you think that Susanna is a reliable source of information about these things, you will at least consider leaving the bus shelter. (In so doing this you are implicitly making three judgments of reliability: of Susanna's perception or her general belief-formation, of your informant's attribution of a belief to her, and of your informant's communication of the attribution to you. These three are regular companions.)

The effect, then, of thinking in externalist terms is to shift the emphasis from the psychology of the individual believer to the reliance that other people place on her beliefs. It pushes us towards understanding "S believes that p" as "to the extent that a is reliable on the topic, others can act on the assumption that p". And this is only marginally a statement about S's psychology at all. In fact, since the reasonableness of relying on a report of S's belief depends on one's relation to the reporter and what other information one possesses, it becomes easier to say what belief-ascription is than what belief is. To ascribe a belief in p to S is to advise or sanction acting as if S were true, if S is a reliable source of information about topics such as p.

Saying this is enough to make a case for what I called above my negative claim: when we abandon internalism functionalism looks very dubious. This raises new questions, though. On the linguistic side it leaves unexplained the conditions under which a sentence uttered by one person about another ascribes a belief in a particular proposition to a person. (I point to a man, whose name you and I know to be "Frederico" and say to you "Susanna believes that his grandmother is a criminal". Which of the following beliefs have I ascribed: "that man's grandmother is a criminal", "Frederico's grandmother is a criminal", "x's grandmother is a criminal" – where "x" is some other name or description of Frederico, "y is a criminal" – where "y" is some other name or description of Frederico's grandmother. There must be pragmatic principles that usually eliminate all but a few of the possibilities; but it is not at all easy to state them.[6]) And on the psychological or metaphysical side it leaves unexplained what properties S has to have for "S believes that p" to be true. In particular, it does not make a connection between stating that we can act as if p, to the extent that S is reliable, and supposing that S has some representation in mind of the proposition p.

---

[6] In this connection see Marina Sbisà's "Belief reports: what role for contexts?" in this issue.

In the remainder of this paper I shall discuss this second question. How do we get from reliability to representations? I shall argue for my central positive claim, that externalism can found an alternative to functionalism. What I say will be incomplete and inconclusive. I believe that the question is hard and deep, and should not be obscured with facile pretended solutions. I begin with some observations about practices of belief attribution.

Real attributive practice So how do we do it? There are two sides to everyday attribution of belief, and though there must be fundamental links between them it is far from clear how in actual human practice the two are connected. The two are, first, our gathering and keeping track of the information available to others, and, second, our reporting this information in the form of assertions of the form "person a believes that s".

There is remarkably little firm psychological data about how we gather and organize our information about the thoughts, memories, and knowledge of others. There is evidence that we categorize people in various subtle ways and make defeasible assumptions about what kinds of information a person of a certain kind will have. (A software engineer from California will know about the state of the San Andreas fault and that of Steve Jobs' career. He or she will understand the electrical supply situation on the west coast, and is likely to have opinions about Boeing's proposed move, about bilingual education, and about the advantages of keeping or exercising stock options. Our assumptions that knowledge of these kinds is present operate so automatically that we rarely notice them[7].) It is clear moreover that we glean a lot of information about the beliefs of those with whom we converse, well beyond the content of the assertions they make. We make use of presuppositions and implicatures, and much less tangible indications contained in a person's choice of words and the general conversational direction. The last factors overlap with another large, obvious, and in detail mysterious source of information: our hypotheses about the reasons for one another's actions. We constantly update our information about the preferences and information of others, in accordance with their conformity or violation of our expectations about them. *How* we do this - what kinds of hypotheses we make, how we generate expectations, how we check expectations against behavior - is extremely mysterious[8].

Mysterious as most of the details are, the following seem fairly safe

---

[7] For data relevant to this see Clark 1992, especially chapters 1 and 2.

[8] For contemporary philosophical accounts of these matters see Lennon 1990. For relevant psychological work, concentrating on the special case of attributing dispositions to future behavior, see Nisbett and Ross 1991. Morton 2002b emphasizes the social aspects of the explanatory hypotheses we form about one another.

assertions. (They are hypotheses, though: most of the evidence is not yet in, and serious philosophy of belief-attribution is only beginning.)

We navigate by knowledge. In keeping track of the information in others' possession, we are most comfortable with true information connected in a non-accidental way with the facts. In other words, knowledge. That is what we should expect, for our information about people is largely about what they are capable of, what facts they have to track in order to be the kinds of people they are, and what sources of knowledge are available to them. The distribution of knowledge follows predictable patterns, while the distribution of conjecture or error is much more individual and unpredictable.

We explain by environmental links. When we explain an action by reference to information in the agent's possession, we trace the agent's path towards a goal along complex routes and around obstacles. We thus rely primarily on knowledge and ignorance: knowledge of effective means to ends and ignorance of obstacles and traps. That is, we start with helpful and obstructive facts and predict that agents will act in accordance with the helpful facts that they know about and run up against the consequences of the obstructive ones that they do not. The core of folk psychological understanding is at a level more primitive than that of belief, where not-knowing does much of the work of believing what is not.

Inference comes last. There are limits to what one can do with knowledge and ignorance. At some point we need to consider what people can infer from what they know. When we do we engage with false belief and with intensionality. We have to take account of the fact that beliefs can be about things that do not exist and can characterize things in ways that are peculiar to the person concerned. Sometimes there are routine bases for this. We know that people are subject to some standard fallacies in reasoning. We know that some kinds of people rely on some systematically misleading information. But this is more often last-resort or show-off stuff. We do it only when we have no safer alternative, when our grasp of another person is extremely secure, or when the main point is not accuracy but entertainment.

Consider now the second linked aspect, attribution proper. When we have gathered information about others' information we can communicate it. Though there is a standard way of doing this, by producing well formed sentences of a language intelligible to attributor and audience, the ways of using this device are varied and subtle. Suppose that the attributor says "Susanna believes that my grandmother is the deepest philosopher since Aristotle". It does not follow that Susanna would say the words "my grandmother is the deepest philosopher since Aristotle", or any near variant of them, since Susanna may not speak English. It does not follow that Susanna's thought describes some person as the speaker's grandmother, since Susanna may not know that the person in question has any relation to the speaker. It does not follow that Susanna's thought makes a

comparison between any person and Aristotle, since all that may be meant is that a very high degree of philosophical depth is ascribed by Susanna to the person (a depth that is extensionally the same as would result from the comparison.) On the other hand it does not follow that these are *not* to be taken as part of the ascription. In some circumstances it will be clear to the audience that some or all of them are intended to be included. The hearer has a delicate job, one that she has learned while learning the rules of conversation and the explanatory practices of folk psychology[9].

Here too our ignorance is very deep and basic. And neither psycholinguistics nor linguistic pragmatics can give a lot of help, for this is a very rich and complex phenomenon in which we are typically drawing on our full range of capacities to understand others rather than on procedures specific to the comprehension of language. But again there are some not very controversial conjectures that can be made without begging some of the harder questions.

The core of attribution is extensional. The most basic aspect of the belief about someone's philosophical skill that is attributed to Susanna is that there is a particular person to whom Susanna ascribes a certain property. Both the person and the property may be further characterized, but once the audience knows who the person is and what the property is they know most of what they need to in order to make use of the attribution in predicting, understanding, or classifying the believer.

Attributions are based on the ontology of attributor and audience. In saying what the object and property are, attributor and audience need to describe them. They do this by relating object and property to the world as they both understand it. They do not normally create a belief-world for Susanna, equipped with exotic objects and properties whose existence she alone assumes. When we do describe someone's belief in terms of some idiosyncratic private ontology, we proceed with great caution, and signal our improvisations and our metaphors.

When the believer's concepts differ from those of the attributor, attribution is improvisatory. Suppose that you are describing the opinions of someone from an alien culture, or someone whose deluded and idiosyncratic inferences have led her to postulate objects and properties that have occurred to no one else, or someone who has so misunderstood some standard vocabulary that she has been led to her own deviant construal of it. If you phrase your attributions in the standard "she thinks that s" idiom you are inviting misunderstanding. The audience will naturally suppose that you are representing her beliefs as referring via your concepts to the objects and properties referred to by your words - by the sentence "s" in some language common to you and your audience, as understood by you and your

---

[9] I am here aligning myself with a tradition that began with Evans 1982, was crucially shaped by Salmon 1986, and has been continued in, among others, Recanati 1991, Crimmins 1992, and Braun 1998.

audience. To prevent this, your attribution will have to be surrounded by special warnings, analogies, explanations of the person's special vocabulary, and so on. (It will look like a commentary on an ancient philosopher.) There is no standard way of doing this. Everyone has to pay attention and find the right mixture of tolerance and criticality.

These three not very controversial hypotheses about the pragmatics of attribution combine well with the three not very controversial hypotheses earlier in this section concerning our ways of gathering information about other's information. They are mutually supportive. If the main point of attribution is to connect the believer to aspects of the environment about which she is likely to be knowledgeable, then the verbal expression of the attribution should aim at describing these aspects rather than the individual details of their representation. If we rarely make straightforward attributions in cases where this extensional focus is inappropriate, then we will reserve attention to an individual's inferential patterns to cases where some conceptual or ontological barrier forces us not to ignore them. If what we know about others concerns most of all their knowledge and their reliability, then when we express what we know our priority will be to specify what the objects of their thinking are.

Two possibilities   These are pieces of a puzzle. They are not all the pieces, and even if they were there might be several ways of putting them together. But two general patterns seem to emerge. To see their appeal, consider two extreme positions, the ones we would get if we focused on one of the patterns and ignored the other.

Knowledge first. The idea here is that our most fundamental understanding is not of belief but of knowledge. We understand when a state of a person is linked to a fact by a process that tracks that fact: if the fact had been different then the state would have been different in some corresponding way. (That can only be a first approximation, but it will do as that.) Then we understand belief as attempted knowledge. We understand for example how someone can know facts around her by perceiving them or by using some other reliable process, and then infer from these facts to beliefs that extend or explain what she knows, and which may or may not be true. This is in effect Timothy Williamson's position[10] (though he intends it as a purely epistemological position, and these remarks construe it a claim about the philosophy of mind.)

---

[10]   Tracking accounts of knowledge go back to Part 2 of Nozick 1981. The "knowledge first" slogan is Timothy Williamson's, see Williamson 2000. Williamson's claims are subtler than those I am making here, with fewer claims about how we actually grasp the concept of belief, and involve less commitment to understanding knowledge in terms of tracking.

From this point of view it is obvious why we focus on knowledge, why we focus on environmental links, and why inference comes last. For in order to attribute a belief one has to first find the knowledge of the environment on which it is based and then trace the inferential or other connection from this knowledge. It is also reasonably clear why attributions are based on the ontology of the attributor and the audience, since knowledge consists in tracking aspects of the real intersubjective world, and in the attributive context it is the ontology of the attributor and audience that is used to define that world. And the improvisational nature of attribution when concepts differ follows for the same reasons. Extensionality, too, falls into place, perhaps a little less convincingly than the others. When a reliable process tracks an aspect of the environment is does so however that aspect is described. If you describe S as knowing by seeing it that a bird is on the branch then you are describing her as using her visual system to keep track of the location of a particular physical object; the tracking works exactly the same way if the object is described as a bird or a small moving thing or aunt Miranda's pet budgerigar. So we can say that an item of knowledge *is* knowledge, without committing ourselves to the details of how the knower represents it. Then if belief is what we derive in various ways from what we know, the same should hold of belief. We should be able to give the core information about someone's beliefs without having to specify the person's full characterization of the objects of belief.

In a nutshell: from this point of view when someone knows that p she represents the fact that p by virtue of the reliable processes that link her to p. The picture of pictures in the head is an image of states of mind being sensitive to aspects of the environment. And when someone merely believes that p she is in a state of mind that is linked to the environment by its derivation from knowledge, and thus can be taken as representing possible facts which under suitable circumstances in which the derivation had led to a true belief would have been known.

attribution first   There is another way of putting the pieces together, while remaining within the externalist point of view. Focus not on belief but on attributions of belief. When Alfonso says that Barbara believes there is a bird on the branch, Alfonso is reporting that there is a bird on the branch, on Barbara's behalf. And this is true if Barbara would herself report that there is a bird on the branch. This too is a first approximation that needs a lot of refining: if Barbara is convinced that speaking of birds portends death then "Barbara believes that there is a bird on the branch" can be true even though she would not herself report it. Assume that such problems can be overcome, and that we can put together an account of belief which takes belief-ascription to be the primary business. This is in

effect Christopher Gauker's position[11]. From this point of view it is not too hard to explain why attributions are based on the ontology of attributor and audience and why attribution is improvisatory when this does not fit the thinking of the believer. For to attribute is to report (on someone else's behalf) and there is no point reporting in terms that do not fit the discussion at hand. On the other hand, from this point of view it is not at all clear why we should focus on knowledge or explain by environmental links. Suppose that Barbara is psychotic, and if asked for her opinion would tell us not only that there is a bird on a branch but also that angels are coming to sing us all to death. Then what we would report on Barbara's behalf is "angels are coming to sing us all to death". So that is what we would truly attribute to her, so what she believes. But it is not knowledge, and not linked to the environment.

There is a definite conflict between the two points of view here. From the 'knowledge first' point of view Barbara's potential assertion about the angels is at best a very dubious and marginal case of belief, while from the 'ascription first' point of view it is a perfectly possible and natural one. The conflict should not be exaggerated, though. The 'attribution first' point of view agrees that attributions of belief that are wildly disparate from the believer's knowledge are in a way irrelevant to the purpose of belief attribution. For the primary point of reporting a fact or opinion on someone's behalf is to get the effect of their observations and thinking, in order to have more reliable information – more knowledge – on which to base our own actions. (The purpose is of a piece with the biological reason for gathering in a herd or a school: there are more eyes, ears, and noses to detect danger. But human gathering is mediated by language; instead of stampeding we say "Barbara thinks there is a wolf over there.") So attributions that present information that is not even a candidate for knowledge are not going to serve the main purpose of attributing. If most attributions were like this the practice would never have become established. But there is no reason why the attribution-first attitude cannot allow other purposes to be served by a practice of reporting for others. Making fun of people or debating their sanity, for example. So the two points of view disagree not about what is practically necessary and likely to be dominate, but on what is possible as an outlying or parasitic case.

Which point of view is right? Do we go with Williamson or with Gauker? I simply do not know. The truth may be that our concept of belief combines both points of view (though there could be concepts of belief which were purely knowledge-centered or purely attribution-centered.) Elsewhere I have tried to defend an account of belief that is in effect a core of Williamson fine tuned with

---

[11]   See Gauker 2002, and Gauker's paper in this issue. An anticipation of the position is found in chapter 5 of Ryle 1949.

Gauker[12]. But I am not at all confident that this is right. What does seem to me incontestable is that when we abandon indefensible epistemological views we have to revise our accounts of what it is to believe and what it is to attribute beliefs. Any revised accounts that incorporate an externalist epistemology will have to build in some way on what is common to the knowledge-first and attribution-first points of view.

References

Braun, David (1998) 'Understanding belief reports', *Philosophical Review* 107, 1998, pp. 555-596.

Clark, Herbert H (1992) *Arenas of Language Use.* Chicago: Chicago University Press.

Crimmins, Mark (1992) *Talk about beliefs*. Cambridge MA:MIT Press.

Dancy, Jonathan (1985) *An introduction to contemporary epistemology.* Oxford UK: Blackwell 1985.

Davidson, Donald (1984) 'Belief and the basis of meaning' pp 141-154 of *Inquiries into truth and interpretation*. Oxford UK: Oxford University Press.

Evans, Gareth (1982) *The varieties of reference.* Oxford UK: Oxford University Press

Foley, Richard (1993) *Working without a net.* New York, Oxford University Press.

Gauker, Christopher (2002) *Words without Meaning* Cambridge, MA: MIT Press.

Goldman, Alvin (1999) 'Internalism exposed'. *Journal of philosophy* 6, 1999, pp. 271-93.

Greco, John and Ernest Sosa, eds. (1999) *The Blackwell Guide to epistemology.* Oxford UK: Blackwell.

Harman, Gilbert (1986) *Change in view.* Cambridge MA: MIT Press.

Lennon, Kathleen (1990) *Explaining human action.* London, UK: Duckworth, 1990.

Morton, Adam (2000) 'Saving epistemology from the epistemologists: recent work in the theory of knowledge'. *British Journal for the Philosophy of Science*, **51,** 685-704

---

[12]   In chapter three of Morton 2002b.

Morton, Adam (2002a) *A guide through the theory of knowledge,* 3'd edition. Oxford UK: Blackwell.

Morton, Adam (2002b) *The importance of being understood: folk psychology as ethics.* London: Routledge.

Nisbett, Richard and Lee Ross (1991) *The person and the situation.* New York: McGraw-Hill.

Nozick, Robert (1981) *Philosophical Explanations.* Oxford, UK: Oxford University Press.

Recanati, Francois (1991) *Direct Reference.* Oxford UK: Blackwell.

Ryle, Gilbert (1949) *The concept of mind.* London: Hutchinson.

Salmon, Nathan (1986) *Frege's Puzzle.* Cambridge, MA: MIT Press.

Williamson, Timothy (2000) *Knowledge and its limits.* Oxford UK: Oxford University Press.