

ELIZABETH FRICKER

AGAINST GULLIBILITY

1.

One main school in the Indian classical tradition of philosophy insists that testimony – ‘learning from words’ – is a source or type of knowledge *sui generis*, one which cannot be reduced to any other type – not to perception, memory, or inference nor, we may add, to combinations of these. Such an irreducibility thesis could take diverse specific forms. One form it may take is as the thesis that a hearer has a presumptive epistemic right to trust an arbitrary speaker. We may essay an initial formulation of this thesis thus:

PR thesis: On any occasion of testimony, the hearer has the epistemic right to assume, without evidence, that the speaker is trustworthy, i.e. that what she says will be true, unless there are special circumstances which defeat this presumption. (Thus she has the epistemic right to believe the speaker’s assertion, unless such defeating conditions obtain.)

The claim that there is such a special presumptive right (PR) to trust associated with testimony constitutes a kind of irreducibility thesis, since the hearer’s right to believe what she is told, on this view, stems from a special normative epistemic principle pertaining to testimony, and is not a piece of common-or-garden inductively based empirical inference.

Testimony’s alleged status as a special source of knowledge is underlined if this PR thesis is conjoined with a negative claim, which we may formulate initially thus:

NC: It is not, generally speaking, possible for a hearer to obtain independent confirmation that a given speaker is trustworthy – that what she says will be true.

If this Negative Claim is true, then knowledge can regularly be gained through testimony only if there is no need for independent confirma-

tion of the trustworthiness of speakers; that is, if the PR thesis holds. So the existence of this special normative epistemic principle is then essential to the gaining of knowledge through testimony. This pair of claims together is one apt explication of the irreducibility thesis of the Nyaya school of Indian philosophy.¹

In this paper I shall give one half of a refutation of the PR thesis, by arguing against the Negative Claim, which features as a premise in one central argument for it. My discussion also shows the *prima facie* case against a PR. A fuller treatment would also consider, and reject, various positive arguments for a PR which may be made, which appeal to the essential nature of language, and of understanding, arguing that these imply that a general disposition to trust is essential to language, and thence to its epistemic legitimacy. Here I can only record my view that no such argument succeeds.

The Negative Claim that there can, generally speaking, be no non-circular confirmation that a given speaker is trustworthy, is false. And any fully competent participant in the social institution of a natural language simply knows too much about the characteristic role of the speaker, and the possible gaps which may open up between a speaker's making an assertion, and what she asserts being so, to want to form beliefs in accordance with the policy a PR allows. The PR thesis is an epistemic charter for the gullible and indiscriminating. This paper argues against gullibility.

2.

To say that testimony is a special source, or yields a special kind, of knowledge, could mean many things. I shall not here take it to mean that testimony constitutes an exception to an otherwise fully general, over-arching conception of knowledge. I take it that its showing knowledge to be, at some level of description, one kind of thing, albeit acquired in different ways, is an adequacy condition on an account of the concept. Such an overarching conception might be causalist or reliabilist. But I favour a justificationist conception, on which a subject's being able to defend her belief appropriately is a necessary condition for it to be knowledge.² The claim that testimony is an irreducible source of knowledge will not then emerge as a counter-example to the thesis that knowledge requires appropriate justification, but as a claim about what kind of justification is required for a testimony belief.³

The PR thesis is such a claim. It is a normative epistemic principle, amounting to the thesis that a hearer has the epistemic right to believe what she observes an arbitrary speaker to assert, just on the ground that it has been asserted: she need not attempt any assessment of the likelihood that this speaker's assertions about their subject matter will be true, nor modify her disposition to believe according to such an assessment. A corollary of the PR thesis is thus that a hearer gives a fully adequate justification of her belief just by citing the fact that "Someone told me so". This simple defence does not need supplementation with evidence for the trustworthiness of her informant. Nor, on this view, does an ordinary hearer need to supplement the simple defence by invoking the PR thesis itself. That thesis is formulated by the philosopher, as a theoretical registering of the fact that the simple defence is all that is needed.

The PR thesis is not to be confused with a descriptive premise that 'speakers mainly tell the truth.' The view that belief in what is asserted is justified by reference to such a descriptive premiss, cited as part of the first-level justification of the belief, is a quite different view, one which would constitute a reduction of knowledge from testimony to an ordinary case of inductively based inferential knowledge. The alleged descriptive premiss (whether claimed to be empirically confirmed fact, or a priori conceptual truth about language) might be invoked in an attempted philosophical argument for the PR thesis. But this is entirely different from its featuring among the premisses which an ordinary hearer must know and be able to cite, to justify her belief.

Our target is the PR thesis. Arguments for it fall into two kinds: the positive arguments from the essential nature of language already mentioned, and a negative argument. This last is a transcendental argument which runs thus:

(1) Knowledge can be and frequently is gained by means of testimony;

(2) [NC] It is not, generally speaking, possible for a hearer to obtain independent confirmation that a given speaker is trustworthy; therefore

(3) There is knowledge gained by testimony only if there is a presumptive right on the part of any hearer to trust an arbitrary speaker; therefore

(4) There is such a presumptive right to trust.⁴

One might reject this argument by rejecting its initial premiss. This is not my strategy. I agree with the proponent of the argument that it is

a constraint on any epistemology of testimony, that it preserve our commonsense view that knowledge can be gained through testimony. This paper is devoted to stopping the transcendental argument by showing its second premise, the Negative Claim, to be false.

3.

The epistemological ‘problem of justifying belief through testimony’ is the problem of showing how it can be the case that a hearer on a particular occasion has the epistemic right to believe what she is told – to believe a particular speaker’s assertion. If an account showing that and how this is possible is given, then the epistemological problem of testimony has been solved.

The solution can take either of two routes. It may be shown that the required step – from ‘*S* asserted that *P*’⁵ to ‘*P*’ – can be made as a piece of inference involving only familiar deductive and inductive principles, applied to empirically established premisses. Alternatively, it may be argued that the step is legitimised as the exercise of a special presumptive epistemic right to trust, not dependent on evidence.

The Negative Claim, when appropriately glossed, is equivalent to the thesis that the first, *reductionist*, route to justifying testimony is closed. The gloss in question is to fix the notion of a speaker’s ‘trustworthiness’ programmatically, as precisely that property of a speaker which would, if empirically established, allow the inference (using only standard principles) to the truth of what she has asserted. As we saw above, the *anti-reductionist* about testimony argues from the alleged closedness of the first route, to the conclusion that the second route *must* be open: to the existence of a special presumptive epistemic right to trust.

It is important to be clear that the only genuine epistemological problem is the one stated above. There is no ‘problem of justifying belief through testimony’ over and above the task of showing that particular instances of testimony can be such as to be justifiedly believed.⁶ The anti-reductionist’s case, I shall show, gains most of its plausibility from confusion over just what the problem to be solved is.

Before we can consider whether the ‘trustworthiness’ of particular speakers can be non-circularly confirmed, and so whether the reductive route to justifying testimony is open, we need to determine just what this property is best taken to be. The first requirement on an

explication of this notion is that it serve the purpose in hand: it must be a property of the speaker *S* knowledge of which suffices, for a hearer *H* on an occasion *O*, to bridge the logical and epistemic gap between ‘*S* asserted that *P*’, and ‘*P*’.⁷ That is to say, if *H* knows that *S* asserted that *P* on *O*, and she also knows that *S* is ‘trustworthy’ on *O*, then she has a basis justifiedly to believe that *P*. Equally (subject to a desideratum explained below), ‘trustworthiness’ should be no stronger than whatever property of *S* it takes to bridge this gap, on particular occasions. If *H* can know that *S* possesses this weakest gap-bridging property on an occasion *O*, this is enough to justify her in believing that *S* asserts on *O*; thus it is only this weakest gap-bridging property which must admit of non-circular confirmation, to provide a reductive solution to the problem of justifying testimony, as we have conceived it. We may also hope that our explication will answer to the intuitive notion of ‘trustworthiness’ of a speaker. It should do so, since the intuitive notion has to it precisely this flavour of ‘that which warrants belief in the speaker’s testimony on an occasion’.

Precisely what trustworthiness, thus programmatically identified, is best taken to be, is spelled out in §7. But we may note here a second theoretical desideratum on our explication.

We may aspire to give a systematic general account of how knowledge (justified true belief) is gained through testimony; or more strictly: of how a subject’s belief may be justified in virtue of its support from testimony. And this account may be conceived as having the following form: A specification of a set \mathcal{T} of sentence-schemata which *characterise* cases of knowledge through testimony in the sense: A hearer *H* has an adequate basis for a true belief of hers to count as justified, in virtue of its support from a certain speaker’s testimony, just when she has knowledge whose content is given by instances, appropriate to the content of her belief, and her situation, of each member of the set \mathcal{T} .⁸

Clearly, a first component of \mathcal{T} will be:

T_1 : ‘*S* asserted that *P* on *O*’.

That T_1 is a necessary component of the set \mathcal{T} (whose members represent a jointly sufficient condition for justified belief) is the hallmark of \mathcal{T} ’s representing what it is for a subject’s belief to be justified by, *inter alia*, evidence from testimony.

And surely it is the notion of trustworthiness, explicated in accordance

with the constraints suggested above, that will furnish the second premise of the desired characteristic set \mathcal{T} ? This is indeed so, if we gloss what it is for trustworthiness to 'bridge the gap' between T_1 and ' P ' appropriately. But we need to be careful about just what this amounts to.

An account which renders perspicuous what is going on in the acquisition of knowledge through testimony must separate out, in H 's total evidence for ' P ', two different strands: The independent evidence for ' P ' which H already has; and the evidence for ' P ' which H gets, given what she knows about S , from the fact that S has asserted it. Effecting this separation is essential, if we are to be able to model what goes on in a 'Humean collision' – that is, a situation where the prima facie evidence for ' P ' from a trustworthy speaker's testimony clashes with strong evidence from other sources against ' P '.⁹ Now specifying a truly characteristic set \mathcal{T} will indeed achieve this separation. But specifying one is not so easy, because for $\mathcal{T} = (T_1, T_2)$ to be characteristic, it is not sufficient, although we may take it as necessary, that the T_1 we choose be such that T_1 and T_2 together entail ' P '.¹⁰

We want our account to separate the two strands in H 's evidence for ' P '. And this implies a further desideratum on \mathcal{T} : its elements should be *epistemically independent of ' P '*, a notion I define thus: No element T of \mathcal{T} must be such that H can know T to be true in virtue of knowing that P and knowing true the other elements of \mathcal{T} . This means that ' P -plus-the-rest-of- \mathcal{T} ' must not together entail T , nor constitute strong evidence for it.

If \mathcal{T} contains a T which is not epistemically independent of ' P ', then a situation is possible in which H knows that P , and knows that which is specified by all the elements of \mathcal{T} , which is not a situation in which she has knowledge that P through S 's testimony; rather, it is one in which the direction of epistemic dependence is the reverse: not: H knows that P , in virtue of knowing all the elements of \mathcal{T} , but: H knows T in virtue of independently knowing that P , and knowing the rest of \mathcal{T} . Such a \mathcal{T} fails to *characterise* cases of knowledge through testimony.

This desideratum that the elements of \mathcal{T} all be epistemically independent of P further constrains the choice of T_2 .¹¹ It rules out choosing the material conditional 'If S asserts that P on O , then P '. This looks like the right choice if we consider only our first requirement, for it is the weakest premise which one can add to ' S asserted that P on O ', to get a pair which together entail ' P '. But it is ruled out by our second desideratum, because it is itself entailed by ' P ', and so H is in a position

to know it whenever she knows that P .¹² If she also knows that S has asserted that P , then she knows the set \mathcal{T} , on this choice of its elements. But she may have no grounds whatsoever for thinking that the material conditional holds of S , other than her knowledge that ' P ' is true. This is not a situation in which she has a basis to know that P on the strength of S 's testimony. On the contrary, it is one exhibiting the reverse epistemic direction. Of course, a situation is *also* possible in which H knows that the material conditional holds of S *not* through knowing that P , but in virtue of knowing something genuinely about S , the intuitive property of 'trustworthiness'. In such a case, she does have knowledge which is based on S 's testimony. The trouble with choosing the material conditional as T_2 , is that the mere fact that S knows the resulting \mathcal{T} does not reveal which of these situations obtains.

The same is true of ' S asserted truly that P ': it too fails the test of epistemic independence of ' P '. The epistemic direction of knowledge through testimony obtains, when H knows that ' S asserted truly that P ' in virtue of knowing that S asserted that P , and knowing something genuinely *about* S – namely, that S is 'trustworthy'. Here, H has knowledge that P in virtue of S 's testimony to it. The reverse epistemic direction obtains, when she knows that S 's assertion that P was true only because she already knows that P . Here S 's testimony adds no further support to ' P ' for H . In requiring that the elements of \mathcal{T} be epistemically independent of ' P ', our idea is precisely to find a \mathcal{T} such that its identity is in itself enough to ensure that the direction of epistemic dependence is always the first, and not the second – i.e. that \mathcal{T} is a characteristic set.

(' S asserted truly that P ' is not a suitable choice for T_2 for other reasons too: it entails ' P ' by itself, while we want a premise which does so only together with T_1 ; and in fact, predicating truth of S 's assertion is an inessential intermediate step, which we can skip, in identifying H 's shortest inferential route from ' S asserted that P ' to ' P ' – c.f. the proposal eventually adopted below).

In describing the direction of epistemic dependence that we want to isolate, I have just employed as a primitive the intuitive notion of S 's 'trustworthiness' which we are supposed to be explicating. But the notion we are groping towards is not doomed to remain an indispensable primitive. We can draw a useful moral from what is wrong with the material conditional. The trouble, in the first instance, is that it is not epistemically independent of ' P '. But this is a symptom of the fact that any

instance of the predicate-schema 'If _ asserts that P on O , then P ', while it is grammatically predicable of S , does not represent a genuine property of S . This last is an intuitive notion we need not attempt to define here; we need only note that a genuine property of S , unlike the material conditional, will not be something which holds of S in a world, merely in virtue of the fact that ' P ' is true in that world. A hallmark of a genuine property of S , in short, is that (special cases apart) it will be epistemically independent of ' P '. To effect the desired separation of the two strands in H 's evidence for ' P ', we must find, as our explication of 'trustworthiness', such a genuine property of S , one such that whether S possesses the property in a world is a matter of what S herself is like. Special cases apart, when 'trustworthiness' is so explicated, situations in which H knows that P , and knows that S asserted that P , and that S is trustworthy, will be precisely those in which, intuitively, we would judge that H has support for ' P ' from S 's testimony; and, flukes apart. H 's evidence confirming S 's trustworthiness will be disjoint from her evidence confirming ' P '.

To find such a notion: which just suffices, together with ' S asserted that P on O ', to entail ' P '; which constitutes a genuine property of S , hence, flukes and special cases apart, is epistemically independent of ' P '; and which constitutes an explication of the intuitive notion of S 's being trustworthy on an occasion of testimony, is our aspiration. A first approximation is the property of S specified by the subjunctive conditional:

Trus1: 'If S were to assert that P on O , then it would be the case that P .'

This bridges the gap and is, special subject matters apart, epistemically independent of ' P '.¹³ Knowing it to hold of S will, generally speaking, require having knowledge about S herself – her character, circumstances, etc. In fact, as we shall see in §7, the property of S specified by this subjunctive conditional is slightly stronger than the choice for T_2 which best fulfils our requirements. We will see there also that the best explication of S 's trustworthiness makes it relative not just to an occasion and an assertion-content, but to a particular utterance U by S . I shall adopt this relativisation from now on, although it is only in our final explication that it is not idle. It is in any case apt, since it is only with respect to her actual utterance that H needs to know that S is trustworthy.

4.

Our final explication of ‘trustworthiness’, and detailed account of how it can be empirically confirmed by a hearer, occupies §§7, 8. But we have enough, armed with the provisional suggestion Trus1, to make some initial points regarding our central concern: the question whether the trustworthiness of a speaker can sometimes be empirically confirmed, so that the reductionist route from ‘*S* asserted that *P*’ to ‘*P*’ is open. The reductionist must make good the following claim (of which, accordingly, the anti-reductionist’s Negative Claim is to be construed as the denial):

Local Reductionist Claim: It can be the case that,¹⁴ on a particular occasion *O* when a speaker *S* makes an utterance *U* and in doing so asserts that *P* to a hearer *H*, *H* has, or can gain, independent evidence sufficient to warrant her in taking *S* to be trustworthy with respect to *U*.

(Notice that to appeal to one’s independent knowledge of the truth of what is asserted by a speaker’s utterance, as evidence for her trustworthiness with respect to it, is not circular; but neither is it a case of possible knowledge through testimony. As we saw above, for just this reason our preferred explication of *S*’s trustworthiness with respect to *U* will not be such that merely knowing the truth of what is asserted by means of *U* is sufficient to establish it. Nonetheless, many instances of independent confirmation of the truth of what a certain speaker asserts provide inductive grounds to attribute a more general trustworthiness to her, as she builds up a track record of independently confirmed accuracy – see below.)

The reduction here claimed is only ‘local’. That is to say, the claim is only that there can be occasions when a hearer has evidence that the *particular* speaker in question is to be trusted with respect to her *current* utterance, without assuming this very fact. I shall call the question whether this local reductionist claim is true the ‘local question’ about testimony. The conception of the epistemological problem of justifying testimony adopted in §3 implies that a local reduction is all we need aspire to, or hope for. A ‘reductionist’ account of knowledge through testimony, in the context of this approach, means such a local reduction of each instance of knowledge through testimony to broader categories of knowledge, and patterns of inference.

Thus on our conception of the problem, justifying testimony by the reductionist route does not, at least in the first instance, require showing that the blanket generalisation, 'Testimony is generally reliable', (or, more simply, 'Most assertions are true') can be non-circularly empirically established. Such globally independent confirmation of the veracity of testimony would require that a hearer have evidence that *most of what she has ever learned through testimony is true*, where this evidence does not in any way rest on knowledge acquired by her through testimony. The fact that such a *global reduction* is not required for it, is crucial to the local reductionist position I argue for in this paper. For, as I readily agree with the anti-reductionist, there are general reasons, stemming from the essential role of simply-trusted testimony in the causal process by which an infant develops into the possessor of a shared language and conception of the world, why the prospects for a global reduction seem hopeless. So *this* negative claim is correct; but beside the present point. Notice therefore how the plausibility of the transcendental argument evaporates, once we identify just what the relevant Negative Claim is. For then we see how modest are the possibilities of non-circular confirmation which it denies, but which are all that is required, for knowledge through testimony to be possible in the absence of a presumptive right to trust.

True, the local reductionist question would transform itself into the global one, if it were the case that the only way of showing that a given speaker was trustworthy with respect to an utterance, was via appeal to the blanket generalisation. But, I suggest, this is not so. The blanket generalisation is actually neither sufficient nor necessary evidence to justify belief, on a particular occasion, that *this* speaker is trustworthy with respect to *this* utterance of hers, which is what it takes to justify belief in what she has thereby asserted.¹⁵ Even if the generalisation were true, there could be circumstances surrounding particular utterances which rendered the speaker's trustworthiness with respect to them doubtful in spite of it. And typically the grounds, when there are such, for expecting a speaker to be trustworthy with respect to a particular utterance of hers, relate to the circumstances and character of the speaker, and the nature of her subject matter; they do not concern the generality of assertoric utterances at all.

More *prima facie* plausible is the claim that the only ground a hearer could ever have for believing a speaker to be trustworthy with respect

to a particular utterance, would be knowledge on her part that that particular speaker is *generally* trustworthy, at least about that kind of thing. Certainly we very often do, quite reasonably, rely on, or distrust, particular individual's testimony on precisely such grounds. But such generalisations about a particular speaker very often *can* be established non-circularly (which amounts to: without reliance on any testimony from that speaker). One means (though not the only, nor the central one, as we shall see in §8) is the approved Humean fashion, induction from observed constant conjunction – we trust one person's report, because she has built up a track record of accuracy; we distrust another because she has accumulated the opposite. And anyway, knowledge of a speaker's general trustworthiness is not the only possible ground for believing her trustworthy with respect to a particular utterance, nor is it always sufficient: someone may be notoriously inaccurate about many things, but one can still reasonably expect her to be right about such elementary matters as what she had for breakfast, or whether she has a headache, or whether a familiar object is on a table in front of her. Conversely, certain circumstances and subject matters provide grounds to expect a generally trustworthy person to be less than reliable – a matter in which she is emotionally involved; something notoriously tricky; when she has been in deceptive or inadequately informing circumstances.

(Note, however, that the *prima facie* incredibility of what a speaker asserts by an utterance is *not* best treated as evidence against her trustworthiness with respect to it. As explained earlier, we need to separate the evidence for '*P*' stemming from the fact that it has been asserted by a trustworthy speaker, from other evidence for or against '*P*'. Where these conflict, there will ensue a Humean battle between them in the belief-updating processes of a rational hearer. To represent this battle most perspicuously, it is the *ex ante* estimate of the trustworthiness of a speaker that we should take; not one revised downwards in the light of her *prima facie* incredible utterance.)

Anti-reductionism about testimony looks plausible if reductionism is so construed as to involve commitment to the claim that the blanket generalisation can be non-circularly established.¹⁶ But my 'local' reductionist can happily grant that this is impossible. There is no need to show that the blanket generalisation can be non-circularly established, in order to show that a hearer can earn herself the right to trust a speaker on an occasion, without needing the gift of a PR; thereby providing a

reductionist solution to the only epistemological problem of testimony which needs to be solved, viz. the local problem.

There is no space in the present paper to consider the reasons why the project of non-circularly confirming the global generalisation is hopeless, nor to defend my view that this does not undermine the rationale for insisting on justification severally for beliefs acquired through testimony. So I shall simply state my views. My view of the global 'problem' about testimony is that it is not a problem. The project of trying simultaneously to justify all of our beliefs which rest in any way on testimony (or equally, to justify a single testimony-belief, but without appealing to any beliefs based on testimony) is not one that is properly embarked on, and we certainly do not need to seek to found these beliefs as a totality in something else. The desire to show that the blanket generalisation can be non-circularly established is an instance of the foundationalist yearning to provide credentials for our system of beliefs from outside that system, or from a privileged subset of it. In this instance this task would be to hive off the part of our belief-system which rests, *inter alia*, on testimony, and show that it can be 'founded' in the remainder which is not. My insistence that the local question is the only legitimate question about testimony is of a piece with a more general coherentist approach in epistemology. Insofar as the anti-reductionist about testimony is expressing an adherence to coherentism, in opposition to foundationalism, I am with her. But this issue of global reductionism, or foundationalism about testimony, comes apart from the issue I am concerned to address. My issue is the local reductionist question: whether, *within* a subject's coherent system of beliefs and inferential practices (in the gradual dawning of light over which testimony will have played an essential part), beliefs from testimony can be exhibited as justified in virtue of very general patterns of inference and justification; or if a normative epistemic principle special to testimony must be invoked to vindicate them and explain their status as knowledge. The issue whether there is a presumptive right to trust not based on evidence is this internal, coherentist issue.

5.

Is knowledge through testimony a distinctive category of knowledge at all? First note that we may define as our epistemic category, and topic of investigation: coming to know that something is so, through knowing

that a certain speaker has asserted it to be so. This definition is restrictive in two respects. First, as to what comes to be known. This restriction is theoretically apt, since there is clearly nothing systematic and general to be said about the unrestricted topic of 'whatever one may be able to infer, on an occasion, from the fact that someone has made an assertoric utterance with a certain content of that occasion'; while we may, as in the present paper, hope to say something general about the inferential path via which a hearer may come to know that which is the content of an assertion, from the fact that it has been asserted. Second, the definition restricts the means by which knowledge of that which is asserted is gained, to being via knowledge of the content and force of the speech act (which will, normally, be obtained through understanding it). This definition excludes, from counting as knowledge gained through testimony, any knowledge gained by one who takes a 'barometer' approach to a group of creatures: that is, who tries to obtain information about the world, from discovering correlations between the sounds the creatures make, and how things objectively are – but who does not regard the creatures as agents nor categorise their utterances as intelligent speech acts. This exclusion is again theoretically apt, since the mechanism involved in gaining any such knowledge is quite different; but in any case, the possibilities for finding such brute phonetic type/environmental-state correlations are very limited, with regard to a fully sophisticated human language-using practice.

But in one respect our definition is permissive: there is no restriction on the subject matter of the speaker's assertion. The domain of potential knowledge through testimony is, on this conception, that of serious assertions aimed at communication, whatever their subject matter. This is at odds with the ordinary language use of 'testimony', which tends to confine it to eye-witness reports of observable events.

Testimony, defined as just suggested, does indeed constitute a distinctive kind of *epistemic link*. There is a distinctive type of connection, characteristic of testimony, between a state of affairs, and a hearer's coming to believe in its obtaining. This connection runs through another *person*, a speaker – her own original acquisition of the same belief, her other mental states, her subsequent linguistic act, which transmits that belief to the hearer.¹⁷ There being this distinctive type of link between a hearer, and what she comes to believe, in testimony, means that there is a distinctive type of justification associated with testimony, in the sense suggested earlier: we can identify a characteristic justificatory schema

\mathcal{T} . A hearer has knowledge through testimony just when she has knowledge whose content is given by appropriate instances of the elements of \mathcal{T} , and can cite such knowledge, or evidence for it, in defence of her belief. But what there is not, this paper argues, is any new *principle of inference* or other normative epistemic principle involved, which is special to testimony.

This makes the ‘problem of justifying testimony’ unlike the ‘problem of induction’. In the latter, the task is to show the legitimacy of a general *principle* of inference, one which is broadly comparable to the principles of deductive inference in the way in which it validates particular inferences of the form in question. It is therefore appropriate to approach the ‘problem of induction’ at a completely general level. The task is to show that an arbitrary inductive inference is valid, by showing that the principle of inference involved in any such inference is a valid one.¹⁸

Now the anti-reductionist may mistakenly suppose that the task of justifying testimony must be approached by looking for some highly general premise or principle which would serve to justify an *arbitrary* testimony belief. Her error stems from a mistaken assimilation of the form of the problem of justifying testimony to that of justifying induction. An anti-reductionist who makes this mistake will start by investigating whether the blanket generalisation ‘Testimony is generally reliable’ can be non-circularly empirically established, with the idea that this general premise, if established, would suffice to justify an arbitrary testimony belief. Finding that such global independent confirmation of testimony is unattainable, she concludes that testimony-beliefs must instead be justified by a special non-empirical normative epistemic principle.

My local-reductionist approach avoids the initial mistake, and so short-circuits the anti-reductionist’s argument. If what were in question *were* a special normative epistemic principle, concerning testimony as a distinctive and unitary category of knowledge, then it would indeed apply indifferently to an arbitrary piece of testimony, and the task of justifying it would need to be conducted at an abstract general level. (Thus positive arguments *for* a blanket PR must indeed be conducted at that level.) But if there is no special epistemic principle in question, and what is common to all and only instances of knowledge through testimony is just a characteristic kind of belief-producing causal process, then there is no reason why what justifies belief in particular instances of testimony must be some proposition or principle applying to testimony

in general. Instead, what justifies a particular hearer's belief in a particular assertion may be her knowledge of relevant facts about that situation and speaker, which warrant her in trusting him. (These will be, as it were, the *foreground* justifying facts – the ones in virtue of her knowledge of which she has gained *this* piece of knowledge through testimony. And which, as a minimum, we may require her to be able to articulate in its defence, for her belief to qualify as knowledge. Of course these facts can bestow knowledge of trustworthiness, and hence of what is asserted, only on a hearer who is equipped with a suitable background of more general knowledge. The account of §§7, 8 will spell out what this is.)

I suggested above that it was hopeless, but fortunately unnecessary for any legitimate enquiry, for an individual to try for *wholly independent* confirmation of the blanket generalisation that 'Testimony is generally reliable'. But it is only on this foundationalist conception of the project of confirmation that it is impossible. A more limited, non-foundationalist version (in which the enquirer makes no attempt to abrogate all of her existing knowledge which depends on testimony) is a perfectly feasible research project. But I think that looking for generalisations about the reliability or otherwise of testimony, in the inclusive sense of *serious assertions aimed at communication of belief*, as a homogeneous whole, will not be an enlightening project. Illuminating generalisations, if there are any, will be about particular types of testimony, differentiated according to subject matter, or type of speaker, or both. True, there is a belief-producing process characteristic of testimony, and consequently a generic type of justification, as captured in \mathcal{T} . This gives one sense in which it is a distinctive and unitary category of knowledge. But when it comes to the probability of accuracy of speakers' assertions, and what sorts of factors warrant a hearer in trusting a speaker, testimony is not a unitary category. The account of how trustworthiness may be empirically established given in §8 below draws on and develops this idea. One aspect of the disunity is, I shall argue, that while there are certain limited epistemic rights to trust involved in particular types of testimony, there is no *blanket* PR to believe what is asserted without needing evidence of trustworthiness, applicable to serious assertions aimed at communication as a whole, regardless of subject matter and circumstances.

6.

In §8 I shall sketch an account of how the trustworthiness with respect to an utterance of a speaker may be confirmed. The kind of confirmation described is, I maintain, often available, and is sought by a discriminating, justifiedly-believing, hearer. The account adopts the standpoint of our commonsense theory of persons and of the nature of speech acts, according to which it is a contingent matter whether a particular assertoric utterance is true, and the speaker trustworthy; and vindicates, within this framework of commonsense theory, the view that a speaker's trustworthiness with respect to an utterance is an empirically ascertainable matter.

But we need first to clarify further the PR thesis which I am opposing. It has several dimensions of possible variation in strength, which must be spelt out, if we are to see just what is the contrast between it, and the view I shall propose.

The 'presumptive epistemic right' in question is a right to form belief in a certain proposition in a certain situation, without needing to have further evidence, or to make further investigations. But we get a weaker, or a stronger thesis, according to what this proposition is. The strongest PR thesis (that is, the one which demands the least of the hearer!) is one which legitimises *simple trust* as capable of yielding knowledge. A hearer has this attitude to a speaker if and only if she is disposed to form belief in any proposition which the speaker seriously asserts in an utterance whose content she grasps; and she lacks the conceptual capacity even to appreciate the possibility that what the speaker says may be false; that is, she lacks a full grasp of our *common-sense linguistics* (CSL), which contains a conception of the nature of language as a social institution, and of the epistemic link which testimony constitutes, including the nature of the speaker's action, and her typical role. (Simple trust is, plausibly, the condition of children at a certain stage in their development.) A simple truster does not have the conception of the speaker's trustworthiness or lack of it, nor appreciate the need for it, so there is no question of her believing in it. A PR thesis endorsing simple trust thus posits an epistemic right on the part of a hearer to believe what is asserted in an utterance, without further conditions, when she has perceived and grasped the content of that utterance; thus in particular without requiring of the hearer-knower the capacity to conceive the trustworthiness of the speaker. (This cagey formulation is required,

since it is doubtful whether one who lacks a full grasp of CSL, though she may respond to an utterance by forming a belief in what is asserted, can be said to conceive the utterance as an assertion in the full richness of that concept.)

A weaker PR thesis, which requires that the hearer be a master of CSL, and appreciate the need for trustworthiness, posits an epistemic right on the part of a hearer to presume an arbitrary speaker to be trustworthy, without needing to have any evidence for this, or to engage in any assessment of the speaker. This thesis is, in the first instance, a licence to believe in the trustworthiness of the speaker; and only derivatively, in the proposition she asserts.

The first, strongest PR thesis makes sense as a thesis about the conditions under which a subject may acquire knowledge from others' assertions (although of course other, 'external' conditions must be added – at the very least truth of what is asserted); but only as part of a reliabilist account of that concept. It cannot be part of any plausible justificationist account, since a subject cannot defend her belief unless she understands the defence; and, as remarked, even the concept of assertion is not available to one who lacks the rest of CSL – lacks understanding that an assertion is, by the nature of the act, not necessarily true, hence the speaker needs to be trustworthy, etc. A simple truster is not in a position to say, with full understanding, even "Someone told me so".

We can therefore leave behind this strongest PR thesis, and consider further only the PR to assume trustworthiness; which restricts the domain of knowledge through testimony to masters of CSL, full participants in the social institution of language, conceptually equipped to play the speaker's, as well as the hearer's role.¹⁹ The point of this PR being the consequent entitlement to believe what is asserted, it is, of course, the minimal gap-bridging property of trustworthiness of the speaker with respect to her current utterance, which is its immediate object. No epistemic right to assume any generalisation about speakers' trustworthiness is needed. Cf. the local/global distinction drawn earlier. Of course the sense in which a hearer is required by our PR to assume, or believe, the speaker to be trustworthy, is not that she is required consciously to form that belief, or consider the question, whenever she forms a belief in what a speaker asserts; but merely that she appreciates the need for trustworthiness, and is disposed to judge the speaker to be trustworthy (or else to abandon her original belief in what

was asserted), when challenged. Implicit belief in trustworthiness will always be attributable to such a hearer, when she believes an assertion.

Our epistemic right to believe (whether in trustworthiness, or in what is asserted), to be at all plausible, must be only 'presumptive' – that is, it must be defeasible in appropriate circumstances. Several dimensions of variation enter here: as to what these 'defeating conditions' (d.c.s) are, and what the hearer's relation to them must be. How strong an epistemic charter our PR thesis is will depend very much on these details of its specification.

A d.c. is, certainly, a condition which cancels the hearer's epistemic right to believe – in the speaker's trustworthiness or, for the strong PR, in what is asserted. That is to say, when the hearer knows one to obtain, she should not form, at any rate not without further evidence, the 'defeated' belief. This gives us a first aspect of the hearer's required relation to a d.c.. On a reliabilist approach, it could be enough that her disposition to believe is thus cancelled, when she is aware of a d.c. But within a justificationist approach, it must be that this disposition of the hearer stems from her appreciation of how the d.c. 'defeats' this belief. Here, there is again a weaker and a stronger option. A d.c. may defeat a proposition, in the sense that it constitutes strong evidence for the falsity of that proposition. Call these *proposition-defeating* d.c.s. Alternatively, it may merely defeat, i.e. cancel, the right to presume that proposition to be true – being a circumstance which indicates that the proposition *may* not be, or cannot be assumed to be true, rather than being definite evidence for its falsity. Call these *presumption-defeating* d.c.s. Clearly, the proposition-defeating d.c.s with respect to any presumptive belief are a subset of the presumption-defeating d.c.s. So a presumptive right to believe in the trustworthiness of a speaker which is cancelled by anything which throws in doubt the presumption that a speaker is trustworthy, will be much weaker – since much more often defeated – than one which is cancelled only by definite evidence of untrustworthiness.

Similarly, a defeasible right to believe in trustworthiness is a weaker epistemic charter for hearers, than a defeasible right to believe what is asserted – since anything which defeats '*P*' will, ex post, defeat the speaker's trustworthiness with respect to any utterance she makes in which she asserts that *P*; while the converse does not hold. The strongest possible PR – to believe that *P*, just on the ground that it has been asserted that *P*, whenever one does not already possess evidence showing '*P*'

to be false, is indeed an epistemic charter for gullibility! But the weakest one: Where the presumptive right is to assume trustworthiness, and a d.c. is any condition which defeats the presumption, by merely raising a question as to the speaker's likely trustworthiness, is a much more limited affair.

There remains a further dimension of variation in the hearer's required relation to the d.c.s, in whichever sense these are taken. The nub of their being d.c.s, is that when the hearer is aware of one, she should not form the 'defeated' belief. When the d.c.s defeat the proposition that the speaker is trustworthy, she should not form belief in it at all; when they defeat the presumption in favour of trustworthiness, she should not believe in it without further investigation: without first engaging in some assessment of the speaker for trustworthiness. The further dimension of variation which remains is: Is the hearer required to look for, be on the alert for, the presence of such d.c.s (of whichever kind)? We know that, when aware of one, she should withhold belief: but is she in addition required to ensure that whenever a d.c. obtains, she will be aware of it, if it is within her epistemic grasp to be so? Or, if not this first, which is a very onerous requirement, then is she at least required to engage in some search for d.c.s, or to be on the alert for the presence or d.c.s?

In fact, the grid of differences set up by our dimensions of variation exhibits some collapse here. Conditions which defeat the *presumption* in favour of trustworthiness are conditions which switch on a requirement to assess the speaker for trustworthiness, i.e. they switch off the right just to assume this without checking on it, the dispensation from epistemic activity which the right to presume trustworthiness constitutes. But to be obliged to keep a constant look-out for any conditions which would suggest that the speaker may not be trustworthy, is not very different from being obliged to assess the speaker for trustworthiness, simpliciter! Such an attenuated PR is not a PR at all: it is not a dispensation from epistemic activity. If the d.c.s defeat the *proposition* that the speaker is trustworthy, the requirement always to be on the look out for such conditions is somewhat less onerous, but still seems not to be very much weaker than a straightforward requirement to assess the speaker for trustworthiness. The notion of a PR, we may conclude, seems only to make sense when it is interpreted as giving the hearer the right to believe without engaging in epistemic activity; when there is no requirement to be on the alert for d.c.s, of either kind.

These considerations reveal the possibility of an interestingly different kind of thesis, which is not a PR, that is, a dispensation from the requirement to assess the speaker; but is rather a thesis applying within the project of assessment, about how it is properly done. I shall call it a *default-position* thesis. To say that a hearer must withhold belief in a speaker's trustworthiness whenever she is aware of signs revealing untrustworthiness, and that moreover she is obliged to be on the alert for such signs, is tantamount to saying the following: the hearer is obliged, always, to assess the speaker for trustworthiness; but within this exercise, the hypothesis of trustworthiness has special status in that it is the default position – it is to be ascribed, in the absence of positive signs of its opposite. The account given in the §8 of how a speaker's trustworthiness may be assessed by a hearer will posit limited default position precepts in favour of what we will shortly identify as the components of trustworthiness.

Our discussion has shown that a PR thesis which is strong enough to be worthy of the name, while fitting into a justificationist framework, is best formulated thus:

PR: An arbitrary hearer *H* has the epistemic right, on any occasion of testimony *O*, to assume, without any investigation or assessment, of the speaker *S* who on *O* asserts that *P* by making an utterance *U*, that *S* is trustworthy with respect to *U*, unless *H* is aware of a condition *C* which defeats this assumption of trustworthiness – that is, *C* constitutes strong evidence that *S* is not trustworthy with respect to *U*; in which case, *H* should not form belief that *P* on the strength of *S*'s assertion that *P*, and should believe, at least implicitly, that *S* is not trustworthy with respect to *U*.

This PR is still programmatic, in that it does not specify just what circumstances would constitute strong evidence against trustworthiness, and there is scope for broader and narrower interpretation here. But it clearly involves what we have identified as the key element of a PR: the dispensation from the requirement to monitor or assess the speaker for trustworthiness, before believing in it. Thus it may be called a PR to believe *blindly*, or uncritically, since the hearer's critical faculties are not required to be engaged. Notice also that it is a *blanket* PR, entitling the hearer to believe in trustworthiness, hence in what is asserted, on any occasion of testimony, whatever the subject matter may be.

(Assuming only that the nature of the subject matter can never in itself constitute strong *ex ante* evidence against trustworthiness.)

It is such a blanket PR to believe blindly that constitutes an epistemic charter for the gullible, and to which I am opposed. The account of how empirical confirmation of trustworthiness is possible set out in §8 involves a limited presumption in favour of trustworthiness, in the very different sense we have identified: it is, in some circumstances, the default hypothesis *within* the critical task of assessing the speaker for trustworthiness.

7.

The thesis I advocate in opposition to a PR thesis, is that a hearer should always engage in some assessment of the speaker for trustworthiness. To believe what is asserted without doing so is to believe blindly, uncritically. This is gullibility. (Though not the only kind. Believing in trustworthiness too easily, i.e. attempting assessment, but doing it badly, is also being gulled!)

So – to return to our central question – if indeed a properly discriminating hearer always assesses a speaker for trustworthiness, what precisely is this property, and how is an empirically-based estimate of it obtainable?

Our method is to develop an epistemology of testimony, including an account of what a speaker's trustworthiness with respect to an utterance consists in, by appeal to the relevant parts of our commonsense theory of the world. This stance is part of a coherentist approach in epistemology: we criticise our belief-forming methods, and standards of justification, from *within* our existing conceptual scheme, rather than attempting to find some mythical point outside it from which to do so.

Now, CSL tells us that, in the normal case,²⁰ a serious assertoric utterance by a speaker *S* is true just if *S* is sincere, i.e. believes what she knowingly²¹ asserts, and the belief she thereby expresses is true. This breakdown is entailed by the commonsense conception of the nature of a speech act of assertion, and of the link between its occurrence, and the obtaining of the state of affairs asserted to obtain. And commonsense person-theory tells us that it is moreover contingent whether any particular utterance is both sincere, and expresses a true belief: it is inherent in the nature of the link, and the psychology of the human subjects who are speakers, that insincerity and honest error are both

perfectly possible. Indeed, commonsense person theory tells us that false utterances are quite common, especially for some subject matters. (This, we may note, constitutes the *prima facie* case against a blanket PR to assume any assertoric utterance to be true, a *fortiori* against one to assume that the speaker is trustworthy. The case is an application of the epistemic precept: 'If a significant percentage of *F*s are not *G*, one should not infer that *X* is *G*, merely from the fact that it is *F*.' A belief so formed is not epistemically rational, which is to say it is not justified.)

In §3 we gave Trus1 as a rough initial explication of a speaker's trustworthiness with respect to an utterance *U* made on an occasion *O*, by which she asserts that *P*. Trus1 is logically equivalent to the claim: 'If *S* were to assert that *P* on *O*, then her assertion would be true'. We have now seen that the truth of *S*'s utterance breaks down (in the normal case to which we confine ourselves) into the utterance's being sincere, and *S*'s expressed belief being true. This suggests that we may frame a more illuminating definition of a speaker *S*'s trustworthiness with respect to an utterance *U* made on an occasion *O*, by which she asserts that *P*, thus:

Trus2: 'If *S* were to assert that *P* on *O*, then it would be the case that *S*'s assertion is sincere, and that the belief she thereby expresses is true.'

Trus2 fulfils our basic requirement on T_2 , that of entailing '*P*' when combined with T_1 . It is more illuminating than Trus1, since *S*'s sincerity, and whether it is likely that if *S* on *O* believes that *P*, then her belief will be true, are what a hearer may, in the first instance, make an empirically-based assessment of. (It is not equivalent to Trus1, since it does not cover the fluke case of a would-be liar who unknowingly utters a truth.)

But the illumination this breakdown provides also shows that Trus2 (and so also Trus1) gives a definition of trustworthiness which is needlessly strong. To be justified in believing what is asserted by an utterance *U* of a speaker *S* on an occasion *O*, a hearer need not know that *any* utterance with that content by *S* on *O* would be sincere; it is enough that she is able to tell that *S*'s *actual* utterance *U* is so. And this difference of strength is empirically significant. We may take sincerity to be a predicate of utterances, and it is very often precisely a

particular utterance that a hearer *H* is able to tell to be sincere, through sensitivity to such features of its delivery as tone of voice, and manner of the speaker. *H* may be able to tell this about an utterance of a speaker who in fact, and perhaps to *H*'s knowledge, is very often insincere – one of whom the stronger sincerity condition contained in *Trus2* is false.

Thus, I suggest, our best and final definition of a speaker's trustworthiness with respect to an utterance *U* is as follows:

Trus(S, U): A speaker *S* is trustworthy with respect to an assertoric utterance by her *U*, which is made on an occasion *O*, and by which she asserts that *P*, if and only if

- (i) *U* is sincere, and
- (ii) *S* is *competent with respect to 'P' on O*, where this notion is defined as follows:

If *S* were sincerely to assert that *P* on *O*, then it would be the case that *P*.

In this definition the relativisation to a particular utterance *U* by *S* is not idle.

Trus(S, U) fulfils, as best we can,²² the requirements explained in §3. It combines with *T* to entail '*P*', and there is no weaker alternative which does so, and which is epistemically independent of '*P*'. *S*'s 'competence with respect to *P*' is defined as in (ii), rather than by a strictly weaker material conditional, in order to fulfil the desideratum of epistemic independence of '*P*', which we saw in §3 that a material conditional fails (equally when the requirement of sincerity is inserted in the antecedent).²³ Notice also that it is right to take the antecedent as in (ii), rather than 'If *S* were to believe that *P* on *O* . . .'. The latter would give a condition which is again unnecessarily strong: perhaps it is only the worlds in which *S* believes that *P* sufficiently confidently to *assert* her belief, that are all *P*-worlds.

Trus(S, U) is weaker than the everyday notion of someone's being a trustworthy or reliable informant, since the latter usually refers to a speaker's assertions more generally, implying that she is generally sincere, and is competent with respect to most of the things she makes claims about. But a person *S* who is untrustworthy, in this generalised sense, can still be *Trus(S, U)*, and known by a hearer *H* to be so, with respect to a particular utterance *U*; in which case, *H* has grounds to

believe what is asserted by that utterance. $\text{Trus}(S, U)$ is the minimal gap-bridging property which we set out to find. As such, it captures the idea that *that utterance* of the speaker is to be trusted.

8.

We have identified the question how a speaker S 's trustworthiness regarding an utterance U may be empirically confirmed as the question how $\text{Trus}(S, U)$ may be confirmed, that is to say, how the sincerity of U , and S 's competence with respect to the content of U , may be confirmed. Notice that these claims are not esoteric, nor technical, but are mere spellings out of what sheer common sense about language, and speakers, tells us.²⁴ Thus in requiring that hearers appreciate the need for trustworthiness, and assess the speaker for it, we are requiring nothing more than what any full participant in the institution of human language is well equipped to appreciate the need for; and, as I shall now argue, can very often achieve.

In recognising an utterance by a speaker as a speech act of serious assertion, with a certain content, a hearer is ipso facto engaging in a minimal piece of *interpretation* of the speaker – ascribing to her an intentional action of a certain kind, and hence at the very least supposing the existence of some configuration of beliefs and desires which explain that action. The theme of my account is: the epistemically responsible hearer will do a bit more of the same. She will assess the speaker for sincerity and competence, by engaging in at least a little more interpretation of her.

A speaker's sincerity and competence, or lack of them, are aspects of her psychology – in the case of competence, in a suitably 'broad' sense, which takes in relevant parts of her environment. Assessment of them is part of, or a prediction from, a more extended psychological theory of her. So, in order to assess a speaker's trustworthiness, a hearer needs to piece together at least a fragment of such a theory of the speaker – an ascription of beliefs, desires, and other mental states and character traits to her. Thus it is commonsense psychology or person-theory, and the related epistemic norms for attribution of these states, that we must look to, to see how trustworthiness can be evaluated.

Notice therefore that while, as we saw in §4, one way of estimating a speaker's trustworthiness is by induction from past assertions of hers independently confirmed as accurate, this is not the best way. As always,

predictions from a *theory* of the subject matter in question – in this case, the psychology of the speaker – will do better, and where there is conflict should override mere extrapolation of observed correlations with no underlying explanation of why they obtain.²⁵

Indeed the primary task for the hearer is to construct enough of a theory of the speaker, and relevant portions of her past and present environment, to *explain* her utterance: to render it comprehensible why she made that assertion, on that occasion. Whether the speaker's assertion is to be trusted will, generally speaking, be fall-out from this theory which explains why she made it; and it is difficult to see how sincerity and competence could be evaluated other than through the construction of such an explanation.

(The need to explain the utterance is sharply felt, when a hitherto reliable informant makes a wildly unlikely claim. – Has she gone crazy? Or been elaborately tricked? Is she kidding? – Or is the best explanation that her outrageous claim to have seen flying saucers is really *true*? We feel at a loss; but it is these alternative explanatory hypotheses that we dither between.)

A psychological interpretation of an individual being an explanatory theory of her, psychological concepts are theoretical in character at least in the respect that their meanings are fixed by their mutual interconnections, and their application to a subject is only holistically constrained by the 'data' to be explained, the subject's actions. Thus the norms which govern ascription of sincerity and competence will be part and parcel of the norms governing the ascription of psychological states more comprehensively. Notice however that norms of ascription – call them *Norms of Interpretation* – whose existence and correctness might be explained by the thesis that they have constitutive status in defining the so-applied psychological concepts, are ones which, at least in the first instance, apply to the highly idealised enterprise of constructing an extensive interpreting description of a person, with 'all' the data to hand; not to the construction of a small fragment of one, on very limited evidence. We shall return to this point below.

I shall first state what I think are the epistemic norms regarding how a speaker's sincerity with respect to an utterance, and competence regarding its content, may properly be estimated by a hearer; and then address the question of why they hold.

In claiming that a hearer is required to assess a speaker for trustworthiness, I do not mean to insist, absurdly, that she is required to

conduct an extensive piece of M15-type 'vetting' of any speaker before she may accept anything he says as true (cf. the implausibly onerous requirement dismissed earlier). My insistence is much weaker: that the hearer should be discriminating in her attitude to the speaker, in that she should be continually evaluating him for trustworthiness throughout their exchange, in the light of the evidence, or cues, available to her. This will be partly a matter of her being disposed to deploy background knowledge which is relevant, partly a matter of her monitoring the speaker for any tell-tale signs revealing likely untrustworthiness. This latter consists in it being true throughout of the hearer that if there were signs of untrustworthiness, she would register them, and respond appropriately.

Such monitoring of speakers, and appropriate doxastic responses formed on its basis are, I suggest, usually found in ordinary hearers, at least to some extent. However, this sort of monitoring for signs of untrustworthiness in a speaker is typically conducted at a non-conscious level. And while its results can generally be fished up into consciousness and expressed, albeit roughly, in words ("I didn't like the look of him"; "Well, she seemed perfectly normal"), no doubt the specific cues in a speaker's behaviour which constitute the informational basis for this judgement will often be registered and processed at an irretrievably sub-personal level. Can a justificationist account of knowledge allow that this kind of process may be knowledge-yielding? Yes, it can: insisting that subjects be able to retail the details of the cues they have responded to is demanding the impossible; but we may insist, compatibly with the sub-personal character of these perceptual or quasi-perceptual capacities, that the subject's beliefs must not be opaque to her, in that she must be able to defend the judgement which is the upshot of this capacity with the knowledge precisely that she indeed has such a capacity – that 'she can tell' about that kind of thing; though she does not know how she does it.

Expert dissimulators amongst us being few, the insincerity of an utterance is very frequently betrayed in the speaker's manner, and so is susceptible of detection by such a quasi-perceptual capacity. But honestly expressed false belief is not so readily detectable, and an informed assessment of a speaker's competence about some subject will typically require that the hearer already know something of the speaker's cognitive talents and failings. How then is knowledge of the latter attainable by a hearer, without, if not an M15-style vetting, then at least

a lot more research than is feasible, when you just want to know the time and have forgotten your watch? As regards sincerity, I suggested that it was tell-tale signs of its absence that a hearer must be disposed to pick up. The flip-side of this coin is that, while there is no right to assume sincerity without monitoring the speaker for it, sincerity is the default position, in assessing a speaker, in the sense we identified earlier; one is justified in taking a speaker to be sincere, unless one observes (and one must be alert for them) symptoms of duplicity.

And, I suggest, the same is true regarding a speaker's competence, *with respect to a certain range of subject matters* – namely, all those for which commonsense person theory tells us that people are nearly always right about such things. Just which topics come within this range is a further question; but it certainly includes such matters as: everyday perceptions of familiar types of item in one's current environment; memories, not too finely specified, of very recent events in one's personal history – such as what one had for breakfast; and a whole range of basic facts about oneself and one's life – one's name, where one works, one's tastes, etc. On such matters, I suggest, competence is the default position – that is to say, one may justifiably assume a sincere assertion by a person of whom one has no previous knowledge to be true, when its subject matter comes within this range, just so long as one remains alert for any sign in their circumstances, or manner, to suggest otherwise, and there are no such signs.

But there are many other possible topics of assertion about which commonsense person theory tells us that people are often, even in some cases usually, wrong. For these subject matters there is no default presumption in favour of competence, and one is not justified in believing what someone says about such things unless one has specific knowledge of their relevant cognitive talents and circumstances.

9.

In virtue of what do these 'default position' norms of attribution in favour of sincerity and, for certain everyday subject matters, competence, obtain? We can identify two opposed views about this. The first view, which is my own, runs as follows: These practical epistemic norms for *ascribing* the psychological attributes of sincerity, and competence, are justified because, and just insofar as, it is fact, and is part of commonsense person theory, that (i) nearly all utterances which seem sincere indeed are so;

and (ii) About these everyday subject matters, where there are no special circumstances, normal people are nearly always right. (Correspondingly, there is no default position in favour of competence for non-everyday subject matters, just because it is not part of commonsense wisdom about persons that they are usually right about these things.)

The opposed view objects to mine as follows: “This explanation gets things the wrong way round. These facts of commonsense person-theory are themselves so as a consequence of the fact that the default positions are epistemic norms governing the ascription of psychological concepts; so they cannot be appealed to to explain or justify these norms. More fully: (i) The obtaining of these norms of ascription guarantees that these ‘commonsense’ facts will be so – so that they are not, as they might seem, contingent, but are features of individuals’ psychology which are guaranteed to be so in virtue of the way psychological concepts are correctly applied. And (ii) the direction of explanation, and justification, is from the existence of the norms of ascription, to the commonsense facts, not vice versa: These norms of ascription are primitive features of psychological concepts, which serve to fix their content; not rules of application which stand in need of justification by appeal to a supposed independently fixed content.”

This opposed view is mistaken, as I shall now show. I think it is plausible that there exist Norms of Interpretation (NIs), in the sense explained earlier: norms for applying psychological concepts²⁶ which have constitutive a priori status, fixing the content of these concepts; so that the truth of an *interpreting description*, as we may call it, of an individual reduces to its fitting the individual in accordance with the correct set of such NIs. But, as mentioned earlier, such a reduction of truth conditions to conditions of ascription will hold, if at all, only with respect to a highly idealised, fancied all-data-in interpretation exercise. And the NIs which apply in such an exercise are by no means the same thing as practical epistemic precepts, applicable in the task of estimating a speaker’s trustworthiness on a very limited basis of evidence about her. Whether they transfer to this limited-evidence (and limited aspiration) case is a further question.

And, I suggest, they do not transfer. It is plausible that ‘Make no unforced attributions of insincerity’, and the parallel principle for false beliefs, are among the correct NIs. But their being so does not ensure that the best interpreting description of an individual will show her as being mainly sincere, or as having mainly true beliefs; that depends on

what departures from the default setting are forced by other NIs. Perhaps there are *also* NIs setting a lower bound on how much insincerity, or false belief, an individual may turn out to have, *salva* the hypothesis that she is indeed a subject of attitudes. But these are further, entirely distinct, constraints. And, I suggest, any such bounds, while being essentially vague, are nonetheless clearly quite low – both for truth of beliefs, and for sincerity of utterances.

If this is right, then it is indeed a contingent empirical fact, not guaranteed by any concept-constituting norms of application of psychological concepts, that, in some given linguistic community, nearly all apparently-sincere utterances are so; and that the speakers in the community nearly always have true beliefs – if not on all subjects (this being palpably false), then at least over a certain quite broad range of subject matters.

There is of course an essentially vague lower bound on the possible incidence of insincerity in a community: beyond a certain point, hearers would cease ever to have the typical responses which are partly constitutive of what it is for a sentence to have a given meaning in a community, and the language would wither away, or change its meaning. But – to reiterate the claim – this lower bound is quite low. In any case, this argument establishes no lower bound on how often any *single* member of a community may lie, *salva* the persistence of language in that community. As regards false belief, I do think it is a priori that for any individual there must be some core range of observable conditions in her immediate environment, such that she is at least disposed to have mainly true beliefs about such matters. If this is not so, she cannot be seen as having the capacity for states of informedness about her environment (which beliefs essentially are) at all. But, once more, this conceptually necessary condition is too weak to affect the current argument.

The ‘default position’ precepts of attribution we have canvassed, applicable in the limited interpretation exercise typically engaged in by a hearer, clearly would not be justified if the commonsense facts which I have suggested to justify them were *not* so; the issue is only as to the direction of explanation between norm of attribution, and commonsense fact. If, as I have claimed, these commonsense facts are not guaranteed to hold by any constitutive attribution-norms for psychological concepts, then their contingent obtaining plays an essential part in the justifying explanation of these default position precepts, and the direction of explanation is as I have suggested: even if there are Norms of

Interpretation, and amongst them default settings in favour of sincerity, and true belief, these do not transfer automatically to the limited-evidence setting, and such limited-evidence default position precepts are justified only by contingent facts of commonsense person theory, and hold only in a community in which these indeed obtain.²⁷

A corollary of my account is that in a community in which these facts which justify the default position precepts were not so, knowledge (though not necessarily belief!) gained from what other people tell one would be much less easily come by, and less widespread. But a language might thrive there nevertheless. Transmission of accurate information is not the only social role and function of the social institution of human language; from many perspectives on human life it is not even the primary one.²⁸

10.

The skeptical reader may want to ask at this point: – Just how different is the proposed account from a PR thesis? And can knowledge of trustworthiness obtained in the manner described really be called empirically based?

For assertions whose subject matter is outside the range for which there is a default position in favour of competence, the contrast between my account and a PR thesis is obvious. But a clear difference remains too in cases in which there is a default position in favour of both components of trustworthiness. My account requires a hearer always to take a critical stance to the speaker, to assess her for trustworthiness; while a true PR thesis, as we have seen, does not. The nub of this distinction is a clear and sharp difference: on my account, but not on a PR thesis, the hearer must always be monitoring the speaker critically. This is a matter of the actual engagement of a counterfactual sensitivity: it is true throughout of the hearer that if there were any signs of untrustworthiness, she would pick them up.

Moreover, as we have seen, the limited default positions in favour of the components of trustworthiness which my account posits, are precepts within the task of constructing a psychological theory of the speaker, not a dispensation from engaging in this task. There is no recognising their defeating conditions except through a general grasp of commonsense psychological concepts, and so the precepts can be conformed to (a fortiori appropriate defence of belief can be given),

only by one who is a master of the latter. Thus, on my account, a person may gain knowledge from others only when she has the needed conceptual framework to conceive and understand them as persons and agents; and moreover engages, at least to some extent, in that interpretative task. The strongest PR thesis we identified earlier does not require this at all; our best formulation, while it required that the utterance is conceived as the speech act it is, did not require any interpretation of the speaker beyond what this itself involves.

Ascribing trustworthiness to a speaker is positing part of a larger psychological theory of her. Such a theory is empirically constrained by, and explanatory of, the speaker's behaviour. The fact that there are certain default settings regarding its construction does not detract from this. In any case the default position precepts do not allow ascription of trustworthiness on no evidence at all: even when trustworthiness is ascribed just on the strength of them, empirical warrant for this is needed, in the sense that the absence of defeaters must have been checked for – as, I have suggested, the hearer will show with such defence as “Well, she seemed perfectly normal”.

But it is important to remember that, as we saw above, while our default position precepts represent what is, given the facts of common-sense psychology, sufficient ground for ascribing trustworthiness to an unknown person, what that person's indeed being trustworthy with respect to her assertion consists in is far from reducing to the obtaining of these limited-evidence ascription conditions. Consequently, while undefeated presumption gives a reasonable basis to believe a speaker to be right about, say, where she lives, one gains stronger confirmation (or disconfirmation!) of her trustworthiness about this and other matters, as one gets to know more about her – acquires more specific knowledge of her relevant cognitive talents and circumstances. A fuller treatment would refine the account offered here by introducing degrees of confirmation, and would introduce into the account of when it is rational (justified) to believe the costs of error: When it matters very much whether what someone says is true, we are less ready to accept what she says without checking her credentials.

11.

We set out to examine whether knowledge from testimony is a special, irreducible type of knowledge. In reviewing what we have discovered,

we may broaden our question to ask not only whether testimony is a special epistemic category, but also whether it is a unitary one. We have found that testimony, appropriately defined, is a distinctive *epistemic link*. That is to say, it is a distinctive type of belief-producing process, and there is consequently a distinctive set of premise-schemata \mathcal{T} recapitulating that process. Appropriate instances of the elements of \mathcal{T} , or evidence for them, when known by a hearer, may be offered by her in defence of a belief acquired through that process, and a belief of hers is known through testimony (*pace* certain qualifications made earlier) just when she is in a position so to defend it.

On the other hand, as regards the likelihood of truth of what is asserted by a speaker, and, consequently, whether a hearer is entitled to presume that she is trustworthy, we have seen that testimony, in the broad sense of serious assertions aimed at communication, is a rag-bag category. This is unsurprising, being a simple consequence of the fact, registered in commonsense person theory, that how likely people are to have true beliefs about a given subject matter depends entirely on what kind of thing it is, and how they are epistemically placed in regard to it. The epistemology of testimony can be no more homogeneous than is the psychology of belief, in this respect.

We have rejected the thesis that there is a blanket presumptive right to trust, applicable to all cases of testimony. Moreover the rag-bag nature of the category in regard of likely truth of what is asserted means that it is a mistake to expect to find *any* epistemic principles as to when one may believe testimony, which apply to all instances of it. Our default position in favour of competence was more selective.

Our account has explained how knowledge may be gained through testimony without recourse to any mysterious epistemic primitives pertaining just to testimony. The limited default positions in favour of sincerity and competence which we have discovered, are epistemic norms within the enterprise of ascribing psychological states to others. Their existence is derived from and explained by the nature of commonsense psychological concepts, whose significance and domain of answerability is much broader than just the explanation of people's assertoric utterances. Thus the conditions under which one may believe another's assertions have been exhibited as fall-out from the nature of commonsense psychological concepts. The epistemology of testimony in this respect is but one part of the broader domain of our knowledge of other minds, and is to be subsumed under that category, not treated as

a separate epistemic domain with its own, irreducible, normative epistemic principles.²⁹

There is another central and fundamental respect in which testimony is a special, and unitary, epistemic category. This paper has taken for granted a hearer's knowledge that a speaker has made a speech act with a certain content and force, and has focussed on the question how she may get from there to knowledge of that which has been asserted. But the epistemology of a hearer's understanding of utterances, and appreciation of them as speech acts, will be at the heart of a full account of how knowledge is gained through testimony. Understanding, whether of one's own or others' utterances, involves special perceptual capacities and kinds of informational states, distinctive of language and of language-using creatures. The epistemology of understanding is intimately bound up with its phenomenology, and with the nature of these special states. Whether or not the best account of how a hearer may know what a speaker has said postulates any normative epistemic principles special to understanding, understanding remains a separate epistemic category in that it involves these special informational states.

The strategy of this paper – to take knowledge of what is asserted as given, and consider the next step – is valid only if the nature of understanding does not itself have implications for that next step. This means, at the very least, that it is not intrinsic to the state of understanding an utterance that it compels the hearer towards belief in what she grasps as being asserted. It is my view that there is nothing in the nature of what it is to understand an utterance which is in tension with the view of knowledge through testimony as inferential knowledge (in the sense that it must be backed by a substantial justification) sketched in this paper, or which provides the materials to defend the presumptive right thesis. But my defence of this claim, and my rejection of other positive arguments for PR, must wait till another day.³⁰

So too must wait further defence of the coherentist epistemic stance within which my account has been developed, from which comes the thesis, essential to my 'local' reductionist approach, that only the local question about testimony needs to be answered, and that it should be answered, as we have done, from *within* the world picture constituted by the 'commonsense' framework of beliefs which we all share; thus that it does not matter, nor does it undermine the rationale of insisting on 'local' reduction and justification, if the global generalisation cannot be independently confirmed by an individual language-user; who will

have made her way into her shared language, and conceptual scheme, through a process in which she was necessarily, at an earlier stage, a simple truster. In this paper I have sought only to block the transcendental argument for a presumptive right thesis, by showing how empirical confirmation of the trustworthiness of a particular speaker is possible.³¹

Magdalen College, Oxford
England

NOTES

¹ I am grateful here to accounts, both written and spoken, of the doctrines of the school, from Arindam Chakrabarti and Bimal Matilal. This pair of theses seems also to be implicit in the anti-reductionist stance of C. A. J. Coady 'Testimony and Observation', *Amer. Phil. Quart.* 10, No. 2, April 1973, pp. 149–55.

² Thus for me, the issue of what it takes for a testimony belief to be justified is one with the issue what it takes for it to be knowledge. Those for whom those issues are not the same – since they favour some other conception of knowledge – may read my account as being simply about justification.

³ I.e. a belief originally acquired through testimony, and whose status as knowledge still rests on that pedigree. In Fricker 'The Epistemology of Testimony', *Proc. Aris. Soc. Suppl.* vol. for 1987, pp. 57–83, I set out a framework which exhibits the complicated interrelations involved here, between original causation, sustaining, and available justifying support of a belief.

⁴ This argument seems to be implicit in Coady *op. cit.*

⁵ In this paper I am assuming that knowledge that such-and-such has been asserted is often had by hearers, and am focusing on the epistemology of the step from there, to knowledge of its truth. See §11.

⁶ If this is shown, then it has been shown that testimony is not just a way of acquiring beliefs, but is moreover one which is capable of yielding knowledge, what we may call an *epistemic link*. Cf. Fricker *op. cit.*

⁷ Throughout my discussion, 'H', 'S', and 'O' are to be regarded as names for an arbitrary hearer, speaker, and occasion respectively. 'P' in contrast must be considered merely a schematic letter holding a place to be occupied by an indicative sentence. Whether outside or inside quotes, 'S', 'H', and 'O', and the possible substitution-instances for which 'P' is schematic, are to be considered expressions of the metalanguage we are using to describe testimony situations. Thus schematic sentences enclosed in quotes, such as 'S asserted that P', constitute (schematic) specifications by us, in our terms, of the content of a hearer's knowledge.

⁸ Instances of \mathcal{T} are sentences of a metalanguage which we use to describe what H knows. There is of course no guarantee a priori that we can thus identify a *single* justificatory schema which covers all and only cases of knowledge through testimony. But it turns out that we can do pretty well. See §5 for how we should define the epistemic link of 'testimony' to this end.

⁹ Cf. David Hume, *An Enquiry Concerning Human Understanding*, Ch. 10. An adequate

treatment of such collisions of contrary evidence would introduce probabilities, as a more detailed model of knowledge through testimony would do throughout.

¹⁰ Of course the grounds justifying a belief need not be so strong as to entail it. The reason for insisting nonetheless that the elements of \mathcal{T} be chosen so as to together entail P , is pragmatic and ad hoc: this represents the best strategy for finding a single characteristic justificatory schema, and the resulting account is illuminating. The possibility of grounds for belief weaker than entailment is allowed for, in this set-up, in the fact that H need only have, and cite, evidence, which may be less than conclusive, that the relevant instances of \mathcal{T} obtain. What may afford this last is endlessly variable, and we cannot hope for a general description circumscribing the possibilities.

¹¹ Note however that it is a desideratum, rather than an absolute constraint, that we thus succeed in characterising knowledge through testimony just by means of our choice of a set \mathcal{T} . Clearly, one cannot find a \mathcal{T} which is epistemically independent of the content of S 's assertion *whatever* the latter may be: c.f., when it is T_2 itself, or evidence for T_2 . But these are special cases, and we may hope to find a \mathcal{T} which is epistemically independent of the content of S 's assertion apart from such cases. As we shall see in §7, it proves difficult to achieve even this perfectly.

¹² Equally, of course, when she knows that S has not asserted that P on O ! But this case need not concern us, since there is no question of H gaining knowledge through S 's testimony, nor of all the elements of \mathcal{T} obtaining.

¹³ An appropriate semantics for this conditional will make it strictly stronger than the material conditional, and with no supposition of falsity of the antecedent. Roughly, it will be true just if all the *nearest* S -asserts-that- P worlds are P -worlds, where the nearness relation is reflexive. It would be nice if a case could be made for a nearness metric which does not have the consequence that the conditional is ensured true whenever ' P ' is a nomological truth. I think the ordinary language locution is rightly heard thus; but finding a regimented semantics with this consequence is another matter. It would, very likely, involve relativising the standard of nearness to the identity of the antecedent.

¹⁴ It is no part of the reductionist position I am arguing for, to claim that empirical warrant for trusting the speaker is available on *every* occasion of testimony. This is clearly false. In cases where it is not, the anti-reductionist and reductionist will disagree over whether the hearer is entitled to trust the speaker, and, in the event she does believe what is asserted, can be said to gain knowledge.

¹⁵ A really strong general claim, to the effect that all, or virtually all assertions are true, *would* suffice to justify belief in an arbitrary assertion, in the absence of further 'defeaters'; and might indeed be employed in a meta-level argument to show the existence of a PR at object level. But a generalisation of this strength is obviously false. (A fortiori is not a conceptual truth about language, as one attempted argument for a PR would claim.)

¹⁶ As I understand it, this is an element in the Indian anti-reductionist case. And Coady *op. cit.* assumes the anti-reductionist must establish generalisations about the reliability of testimony.

¹⁷ Is this connection causal? Its latter stages which are our primary concern always are, but whether the speaker's initial acquisition of her belief can be thought of as caused by its subject matter depends on what kind of thing that is, and how her belief arose.

¹⁸ If the reader is unhappy with this view of the problem of induction, she may consider the justification of deduction instead, which surely takes this form.

¹⁹ Is this unkind to children? The upshot of my casual discussions with developmental

psychologists is that they (children) acquire the ability to lie, and so maybe the grasp of CSL which shows this possibility, remarkably early. But a feeling that my theory is too demanding on hearers may anyway be an intuition against the requirement that knowledge requires justification, rather than against my account of what it takes for a testimony-belief to be justified.

²⁰ Freak cases are possible – where a would-be deceiver happens to have a compensatingly false belief. But for our project, of giving a systematic general account of how knowledge is gained through testimony, we may set these aside, taking the normal case as our domain.

²¹ That *S* understands her own utterance we may consider to be packed into the fact that it is a serious assertion. The epistemology of such knowledge is outside the scope of this paper.

²² Note that the definition of competence given does not allow any inference ‘backwards’ to sincerity, from knowledge of competence and the truth of what is asserted; but a stronger definition – ‘If *S* were to believe that *P*, then it would be the case that *P*’, would do so. Intuitively, this kind of ‘backwards’ knowledge of sincerity can occur. There is another difficulty, viz. that one may also know competence backwards, when ‘*P*’ expresses a necessary truth and one knows this fact; and equally, in the absence of a semantics which avoids this, when one knows it to express a nomological truth (see footnote in §3 above). But there is no alternative which meets our requirements better than the \mathcal{T} consisting of T_1 and $\text{Trus}(S, U)$; so we must perforce complete our characterisation of knowledge through testimony by putting restrictions on how sincerity and competence are known by *H*, which rule out these cases of ‘backward’ confirmation.

²³ The present account thus differs from the one I offered in Fricker *op. cit.* There I opted for a material conditional expressing ‘competence with respect to *P*’, for the prima facie reason in its favour, that it is the weakest further premise which validates the inference to ‘*P*’. I now hold that earlier choice to be wrong because it fails the test of epistemic independence.

²⁴ That it takes some care to arrive at a correct theoretical definition of trustworthiness in no way undermines this claim. The difficulty of formulating explicitly conditions of which we all have a sure implicit grasp, is the general experience with analyses of ordinary concepts.

²⁵ If Russell’s chicken had only interpreted its feeder, her murderous intent on that last day would not have come as such a surprise!

²⁶ And with them, simultaneously, semantic concepts, of course. My discussion here is too brief to bring in explicitly the fact that, in any ascription of psychological states to an individual, the meaning of the sentences she utters are always, at least in principle, also in the melting pot. But nothing I say here is in neglect of this fact, which does not invalidate the argument of this section, in particular the claims that any conceptually-ensured lower bounds on false belief, and false utterance, are quite low.

²⁷ If considerations about interpretation do not suffice on their own to justify a default position in favour of trustworthiness, then a fortiori they do not serve to justify a PR thesis. This is one of the attempted ‘positive arguments’ which, in my view, does not work.

²⁸ My views here have been influenced by discussions with Prof. Mike Gilsenan, about his experiences as an anthropologist studying Middle Eastern societies. There is of course *much* more to be said on these matters.

²⁹ It is itself part of that broader domain, rather than reducing to it, in that, as already noted, semantic and psychological concepts hang together, fitting simultaneously onto a subject.

³⁰ See Fricker *op. cit.*, pp. 74–5.

³¹ I am grateful for comments from Michael Bacharach, John Campbell, Bill Child, Dale Jamieson, Philip Pettit, and Tim Williamson. I am also particularly indebted to Arindam Chakrabarti, whose vigorous defence of the Indian view provoked this paper.