

This is the penultimate version of a paper that is forthcoming in *Studies in History and Philosophy of Science*. For citation, please refer to the final version at the publisher's website.

## **Reconsidering the role of inference to the best explanation in the epistemology of testimony**

Axel Gelfert  
National University of Singapore

### **Abstract**

In his work on the epistemology of testimony, Peter Lipton developed an account of testimonial inference that aimed at descriptive adequacy as well as justificatory sophistication. According to 'testimonial inference to the best explanation' (TIBE), we accept what a speaker tells us because the truth of her claim figures in the best explanation of the fact that she made it. In the present paper, I argue for a modification of this picture. In particular, I argue that IBE plays a dual role in the management and justification of testimony. On the one hand, the coherence and success of our testimony-based projects provides general abductive support for a default stance of testimonial acceptance; on the other hand, we are justified in rejecting specific testimonial claims whenever the best explanation of the instances of testimony we encounter entails, or makes probable, the falsity or unreliability of the testimony in question.

### **Introduction**

In this paper, I reconsider the role of inference to the best explanation as applied to the problem of testimonial knowledge. Testimony is perhaps the most important source of knowledge in our epistemic lives: on reflection, there is very little that we could claim to know entirely off our own bat. Likewise, inference to the best explanation is one of the most powerful cognitive tools at our disposal: it allows us to infer new knowledge and to gain understanding of the things we already take ourselves to know. Combining testimony and inference to the best explanation, then, should be a powerful recipe for understanding how we make sense of the world we are in.

This paper is organised into six sections. The first summarises the contemporary debate in the epistemology of testimony, by focusing on reductionist and anti-reductionist accounts of the justification of our testimony-based beliefs. The second introduces the idea of inference to the best explanation (IBE) and notes some often overlooked early connections

with the debate about testimony. A basic sketch of how IBE can be applied to testimony is discussed in the third section. This is followed by an interlude, in which I critically discuss a recent proposal by Peter Carruthers, which integrates IBE, testimony, and our very capacity to think creatively into an intriguing evolutionary account. The fifth section discusses in detail some of the attractions and problems faced by standard IBE-based accounts of testimony. In the sixth and final section, I develop a novel IBE-based account of testimony as a source of knowledge. In particular, I suggest that IBE plays a dual role in the management and justification of testimony: First, the coherence and success of our testimony-based projects provides general abductive grounds for a stance of default acceptance of testimony; second, when we reject specific instances of testimony – as, on occasion, we must – we are justified in doing so if the best explanation of the testimony we encounter entails, or makes probable, the falsity, or unreliability, of the testimony in question.

### **The epistemology of testimony**

The philosophical debate about the status of beliefs acquired from testimony usually takes place between two schools of thought. (For a survey of the contemporary debate, see Adler 2008.) On the one hand there are those who argue that, all else being equal, we are entitled to accept a piece of testimony ‘as is’, without any requirement to take active steps to determine its veracity. On this account, we have a default entitlement to believe what we are told. This anti-reductionist view is often traced back to Thomas Reid (1710-1796), who believed the exchange of testimony to be governed by principles ‘implanted in our natures’ by ‘[t]he wise and beneficent Author of nature, who intended that we should be social creatures’ (1764: VI: xxiv). More specifically, Reid posited two complementary innate principles: the *Principle of Veracity*, i.e. a propensity to speak the truth, and the *Principle of Credulity*, according to which humans have a propensity to believe what they are told. (Not surprisingly, anti-reductionism has also been referred to, sometimes disparagingly, as *credulism*.) Opposed to this school of thought stands an impressive line-up of reductionists about testimonial warrant. Beginning with David Hume (1711-1776), the idea has been that testimony cannot possibly be a fundamental source of knowledge, since it depends, in an obvious way, on sense data for its reception, and on the competence and sincerity of the speaker for its truthfulness. Hence, whatever justification testimony-based beliefs might possess, must eventually be derived from more basic epistemic sources, such as perception, memory, and inference.

Describing the history of the philosophical debate about testimony as an ongoing controversy between reductionists and anti-reductionists is, of course, a gross oversimplification.<sup>1</sup> What this antagonism brings out, however, is the existence of a set of opposing intuitions regarding our response to what other people tell us. These can be made more explicit by formulating principles concerning the source of epistemic justification,

---

<sup>1</sup> For important alternative traditions in the philosophical thinking about testimony, see for example (Gelfert 2006) and (Jardine 2008).

considered from the perspective of the recipient of testimony ('the hearer'). Anti-reductionism can then be seen to entail a *presumptive right* (PR) thesis:

*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.) (Fricker 1994: 125).

The suggestion is not that the PR thesis somehow guarantees the truth of beliefs formed on its basis: the epistemic right referred to in the PR thesis is at best a *defeasible* one. However, it does establish a default mode of acceptance: all else being equal, and in the absence of contravening evidence, the PR thesis justifies the acceptance of a speaker's testimony without any further investigation. Testimony, from an anti-reductionist perspective grounded in the PR thesis, is a source of justification in its own right.

By contrast, reductionism is based on the idea that any justification a hearer might have for accepting another person's testimony, and any epistemic justification of the belief thus acquired, must derive from more fundamental sources of epistemic justification. Hence, the reductionist thesis (R):

(R) On any occasion of testimony, the hearer's epistemic right to believe what she is told must be grounded in other epistemic resources such as perception, memory and inference.

Much debate has revolved around the question of whether such reduction of testimonial justification to the more basic epistemic sources of perception, memory, and inference, is indeed possible, or whether the reductionist thesis is really tantamount to demanding the impossible. This much seems obvious: a '*pessimistic* reductionist', who subscribes to (R), but does not believe that such reduction is actually achievable – either as a matter of principle, or for all practical intents and purposes – will be forced to give up much of what we ordinarily take ourselves to know. In its most extreme form, pessimistic reductionism will lead to a strong form of scepticism.

Given that virtually all reductionists agree with non-reductionists that testimonial scepticism is no live option, all contributors to the debate operate under what has been called the '*commonsense constraint*', i.e. the acknowledgment that 'testimony is, at least on occasion, a source of knowledge' (Fricker 1995: 394). For reductionists, however, this creates a special burden of proof, since they must now demonstrate just how testimonial knowledge derives its justification from the more basic epistemic sources. Hume famously attempted a '*global*' reduction of testimony, by assimilating our acceptance of testimony to simple enumerative induction: We are justified in relying on testimony not in virtue of 'any connexion, which we perceive a priori, between testimony and reality, but because we are accustomed to find a conformity between them' (X.1: 172). There is general agreement that Humean global reductionism is a hopeless project, due to a number of fatal problems, ranging from the paucity of first-hand evidence of the '*conformity*' in question to reference class problems and a general inability of Humean reductionism to account for the manifest success of our testimonial practices. (See Lipton 1998: 14-21.)

But not all forms of reductionism are necessarily of the Humean ‘global’ variety: *Local* reductionism concedes that testimony is ineliminable as a matter of principle, because it is necessary for the acquisition of a conceptual framework during childhood. As a result, a certain subset of testimony-based beliefs is admitted as irreducibly justified, simply in virtue of our having acquired them at an early developmental stage. Once a stable conceptual framework and proper cognitive capacities are in place, however, the demand for reduction becomes effective. Hence, local reductionism distinguishes between a ‘developmental’ stage and a ‘mature’ phase in our epistemic lives, which gives rise to the local reductionist thesis:

(LR) In cases of a mature recipient gaining knowledge from testimony, the epistemic right of the recipient to believe what he is told can, and must, be grounded in other epistemic resources such as perception, memory and inference.

Unlike the Humean attempt at global reduction, (LR) acknowledges the *de facto* irreducibility of *some* testimony-based beliefs, while affirming the basic reductionist intuition that *mere reliance* on testimony is never quite good enough for *mature* reasoners: ‘We know too much about human nature to want to trust anyone, let alone everyone, uncritically.’ (Fricker 1995: 400). Instead of simply taking on trust what others tell us, local reductionism demands a critical assessment of the speaker’s trustworthiness. In particular, the hearer ‘should be continually evaluating [the speaker] for trustworthiness throughout their exchange, in the light of the evidence, or cues, available to her’; this may require a (passive) readiness ‘to deploy background knowledge’ as well as (active) ‘monitoring [of] the speaker for any tell-tale signs revealing likely untrustworthiness.’ (Fricker 1994: 150). Failure to continuously exercise one’s responsibilities as a mature reasoner, for example by accepting what one is told ‘without any investigation or assessment’, would constitute ‘an epistemic charter for the gullible’ (ibid.).

The positions I have sketched in this section show that there is a whole spectrum of possible stances towards the problem of testimonial knowledge, ranging from a stance of default acceptance to an outright dismissal of testimony as an independent source of knowledge. What makes local reductionism interesting, is its attempt to steer a middle path between those two extremes. While it fails for independent reasons, some of which are due to internal instabilities *within* the position (for a detailed argument, see Gelfert 2009), local reductionism is relevant in the present context, because it has often been described as being based on inference to the best explanation. And indeed, there are elements in Fricker’s local reductionism that suggest such an interpretation, for example when she writes that the hearer ‘must engage in a piece of psychological interpretation of her informant, constructing an explanation of her utterance as an intentional speech act’, and that assessments ‘of her sincerity and her competence, or their lack, will be part of this explanatory mini-theory’. (Fricker 1994: 404). But it is worth pointing out here that local reductionism has a very narrow view of the role of IBE in testimony, limiting it to the ‘mature phase’ only and placing an almost puritanical emphasis on the amount of ‘epistemological work’ we must put in before accepting what others tell us. According to local reductionism, truly responsible testimonial beliefs are only formed later in one’s life. While some testimonial knowledge is acquired during the epistemic agent’s developmental phase, ‘only later, during her “mature

phase,” is it based on an inference to the best explanation’ (Malmgren 2006: 231). In the present paper, I want to argue that the role of IBE in accepting testimony runs much deeper than this and, in particular, is not only compatible with, but *justifies*, a default stance of acceptance in most contexts.

### **IBE, testimony, and evidence: early connections and controversies**

Our capacity for knowledge and understanding of the world has increased dramatically over the course of human biological and cultural evolution. With the advent of collective projects of inquiry, such as science, we are in a position to know many more things than we could ever find out about first-hand, including many items – such as past events, unobservable entities, or hidden causal mechanisms – that we could never, *as a matter of principle*, experience directly. However, in our cognitive projects, we do not merely satisfy our general curiosity for new items of knowledge: we also aim to integrate what we already know into a coherent overall picture, i.e. we aim to increase our understanding. In particular, we have an interest in arriving at explanations of things we already take ourselves to know: we want to know *why*. Children intuitively grasp the significance of that simple interrogative when they realise that any answer to a previous why-question can be followed up by another round of asking ‘Why?’.<sup>2</sup> Likewise, when a scientist carries out a litmus test and infers the presence of acidity from the colour of the litmus paper, he is making an inference to an explanation of *why* it is that the strip of paper turned red (*rather than* blue, or remain neutral).

In his classic 1965 paper, in which he coined the phrase ‘Inference to the Best Explanation’ (henceforth IBE, for short), Gilbert Harman describes the distinctive character of explanatory inference: ‘In making this inference one infers, from the fact that a certain hypothesis would explain the evidence, to the truth of that hypothesis.’ (Harman 1965: 89) Typically, however, there will be more than one hypothesis which might explain the evidence. This complicates considerably the task of inferring explanations: given that the evidence is compatible with a range of competing explanatory hypotheses, the task becomes one of identifying the hypothesis which, if true, would provide a ‘better’ explanation than would any of its competitors. Exactly what makes one hypothesis a significantly better explanation than another hypothesis? Giving a general answer to this question has proven extraordinarily difficult. Harman argues that we evaluate explanatory hypotheses along different dimensions of goodness, such as ‘which hypothesis is simpler, which is more plausible, which explains more, which is less *ad hoc*, and so forth’ (ibid.). Simplicity, plausibility, explanatory power, and the lack of *ad hoc* elements are, of course, familiar criteria of theory choice in the philosophy of science. And indeed, ‘IBE’ is often – though

---

<sup>2</sup> In an endearing autobiographical anecdote, Peter Lipton described the moment ‘it dawned on me that, whatever my mother’s answer to my latest why-question, I could simply retort by asking “Why?” of the answer itself, until even my mother ran out of answers or patience.’ (Lipton 2004: 22) In the slightly more embellished version he used to tell in lectures, the why-regress unfolded as he and his mother were walking down Fifth Avenue along Central Park, where it came to an abrupt end – with a slap on Peter’s cheek – ‘just outside the Guggenheim Museum’.

sometimes hastily – treated as synonymous with C.S. Peirce’s term ‘abduction’, which refers to a third mode of inference, alongside deduction and induction, and which is intimately related to the generation of new hypotheses. The idea of a third, ‘creative’ mode of inference was subsequently developed in debates concerning the ‘logic of discovery’ in science and continues to inform the philosophical debate. (See Paavola 2006.)

When introduced at this level of generality, IBE seems to be little more than a programmatic idea. In the 1991 edition of his monograph *Inference to the Best Explanation* – now a modern classic in the philosophy of science – Peter Lipton motivated the need for a book-length study on the subject by observing that IBE ‘is more a slogan than an articulated philosophical theory’ (Lipton 1991: 2). If one wants to show that IBE is an improvement over competing philosophical accounts of explanation, and that it provides a descriptively more adequate account of our inductive practices, then one needs to spell out the slogan in more substantive ways. There are at least three different ways in which Lipton’s account puts flesh on the skeletal programme of IBE sketched by Harman. First, IBE is tied in with a causal model of contrastive explanation, which is explicitly intended as an alternative to, and an improvement over, the hypothetico-deductive model of explanation. (Lipton 2004: 88) (The importance of contrastive explanation was already hinted at in the litmus-test example above, according to which the why-questions asked by scientists often take the form of ‘why A rather than B?’.) Second, IBE as a general account of explanatory inferences needs to be (partially) freed from the connotations of ‘explanation’ as a success term: IBE cannot be understood as inference to the best of the *actual* explanations, since such a model would make us ‘too good at inference’. If IBE is to be a descriptively adequate account of our inductive practices, it must recognise our fallibility: ‘we sometimes reasonably infer falsehoods’. (Lipton 2004: 57). Finally, once one realises that some *potential* explanations can be deemed good explanations irrespective of whether they are *actual* or not, one can then go on to distinguish two senses in which a potential explanation may be the best among a pool of ‘live explanatory options’: on the one hand, it may be the most probable explanation, on the other hand, it may be the explanation which, if true, would provide the most understanding. In Lipton’s memorable turn of phrase, the choice is between the ‘likeliest’ and the ‘loveliest’ explanation. Whether such an ‘enriched’ account of IBE can live up to the edifying spirit of the initial slogan then becomes a question of whether explanatory power is indeed truth-conducive, and if so under which conditions. This much seems clear: Whether, and when, the loveliest explanation does in fact coincide with the likeliest potential explanation, will depend on the class of cases under consideration.

While it is true that IBE, in its initial formulation, was more of a slogan than an exhaustive account of our inductive and explanatory practices, it is noteworthy that Harman’s original proposal was not entirely abstract, but was motivated by a consideration of concrete examples. On Harman’s account, the archetypes of an ‘end user’ of IBE are *the detective*, who ‘puts the evidence together and decides that it *must* have been the butler’, and *the scientist* who, in inferring the existence of atoms and subatomic particles, ‘is inferring the truth of an explanation for various data which he wishes to account for’ (Harman 1965: 89). But for Harman, the use of IBE extends beyond these obviously inferential activities of the

detective and the scientist, to more mundane cases of everyday belief formation. Significantly, this includes our acceptance of another person's testimony: our confidence in another's testimony 'is based on our conclusion about the most plausible explanation for that testimony'. More specifically, when 'we infer that a witness is telling the truth, our inference goes as follows: (i) we infer that he says what he does because he believes it; (ii) we infer that he believes what he does because he actually did witness the situation which he describes.' (Harman 1965: 89). While one might argue that Harman's conception of testimony as limited to eyewitnesses is overly narrow, nothing much hinges on this restriction and his example already contains the basic elements of an IBE-based account of testimonial knowledge in general, to be discussed in the next section. Why did this early connection between IBE and the problem of testimony go unnoticed, both by proponents of IBE and by those epistemologists who, in the last two decades, rediscovered testimony as a topic worthy of philosophical analysis? One reason might be the particular twist that Harman gave to his account, which was concerned more with demonstrating that *enumerative induction* is warranted only to the extent that it is a special case of IBE, and less with analysing IBE as a general pattern of reasoning and as a powerful tool for assessing different mechanisms of belief formation. Another reason might simply be that, as discussed above in connection with local reductionism, we typically do not experience acceptance of testimony as inferentially very demanding. As Lipton puts it, whereas 'scientists and detectives sometimes have to do considerable work to come up with the explanations they will go on to infer', in the case of mundane testimony 'the belief we come to acquire is given to us on a plate, since it is simply the content of the testimony itself.' (Lipton 1998: 25).

Before considering in detail the prospects of an IBE-based account of testimonial knowledge, let us briefly turn to the evidential base of explanatory inferences, both in general and in relation to testimonial behaviour. While IBE need not be committed to any one conception of evidence in particular, it does demand that the explanations we infer be compatible with our best available evidence. Of course, this way of putting it leaves open precisely what should count as our 'best' evidence, and when it should be deemed 'available'. Mere *in-principle* availability might be too strong a demand if, for whatever contingent impediments, we cannot access the evidence in question. In order to limit the pool of possible explanations to those that are 'live options' – i.e. *potential* explanations, *given* our evidence – it may be best to go along with Lipton's proposal that 'a potential explanation is any account that is logically compatible with all our observations (or almost all of them) and that is a possible explanation of the relevant phenomena.' (Lipton 2004: 59). If it is indeed compatibility with (actual) observations that exhausts the demand for evidence, then the onus is on us to ensure that we have made the relevant observations and are not drawing hasty inferences. This becomes especially pertinent in the case of testimony. Imagine I am told that  $p$  by someone I have no reason to distrust. Assuming that  $p$  is neither implausible for independent reasons, nor self-evident, on what evidence can I come to know that  $p$ ? Being told that  $p$  does not entail that  $p$ , so even if we *eventually* accept what we are told, and go on to treat our belief that  $p$  as evidence for all kinds of things,  $p$  itself cannot plausibly serve as evidence of the truth of the testimony without inviting the charge of begging the question.

But when we are told that  $p$ , we have more to go on than just the content of the testimony: we also have the fact that we have been told that  $p$ . On one plausible interpretation, then, the evidential base that supports our testimonial practices consists of *tellings* – i.e., facts about who said what, and when – rather than of the *contents* of what we are told. (It is perhaps worth noting that such a view does not yet come down on either side of the reductionism/anti-reductionism divide and, in particular, is not committed to reductionism: Tyler Burge, to mention just one prominent anti-reductionist position in the epistemology of testimony, regards the fact *that something has been intelligibly presented to us as true* as itself evidence of the rationality of its source, and hence as issuing in a *prima facie* entitlement to accept what we are told. See Burge 1993: 469.) But one need not stop with the bare recognition of the empirical fact of tellings: each instance of testimony is accompanied by a great variety of circumstances, about which we can make all sorts of observations. It is this richer interpretation of testimonial evidence that Fricker draws on when she demands that we should ‘monitor’ our interlocutors ‘for any tell-tale signs revealing likely untrustworthiness’ (Fricker 1994: 150).

An important earlier influence on the philosophical debate about the evidential basis of testimonial knowledge, which predated many of the controversies discussed in the previous section, was Paul Grice’s account of the nature of the communicative act. (Grice 1957 & 1969). According to what has come to be called the ‘Gricean mechanism’, the speaker  $S$  aims to get her hearer  $H$  to believe something (e.g., that  $p$ ) by getting him to recognise that this is what she is trying to do. The way in which  $H$  is intended to arrive at his belief that  $p$  is by taking  $S$ ’s act of telling (e.g., her perceptible utterance of certain sounds) as evidence of her intention to get  $H$  to believe that  $p$ , and subsequently to take her having that intention as evidence that  $S$  herself has the belief that  $p$  (or, on a somewhat stronger interpretation, takes herself to know that  $p$ ), which in turn is then to be taken as good evidence that  $p$  is indeed true. What  $S$  is relying on when she intends to communicate her belief that  $p$ , is the assumption that  $H$  will take the fact that  $S$  wants him to believe that  $p$  as something which, given proper background beliefs and competencies on the part of  $H$ , is a good reason for  $H$  to believe that  $p$ . Note that, on this interpretation of Grice’s account, the speaker need not offer any direct evidence that would independently support what she claims, nor need she offer any natural evidence as to her own sincerity or competence. Grice equivocates slightly on this point, though. On the one hand, he concedes that ‘from one point of view questions about reasons for believing are questions about evidence’ (Grice 1957: 358), on the other hand he makes clear that no more than recognition of the speaker’s intention to inform is required for the hearer to be justified in bringing himself to believe that  $p$ .

The emphasis, in the Gricean programme, on the speaker and her intentions, to some extent obscures what follows from it for the perspective of the recipient of testimony. Just how much do we need to know about the speaker’s intentions before we can accept her word? Angus Ross has argued that there might be an even more fundamental problem with the evidential view of testimony, which is perhaps best described as akin to a form of ‘cognitive dissonance’ in the speaker. Whereas I can, in my capacity as speaker, see my choice of words

as constrained, in a certain sense, ‘by my obligation to be truthful, helpful, or discrete’ (or by similar considerations), I cannot ‘read my choice of words as evidence for the existence of the state of affairs they report’ without making ‘some assumptions about my nature in the sense of my inclinations, preferences, and commitments’ – assumptions which I cannot see as constraints on my choice of words without a peculiar ‘form of disengagement from my own actions, what Sartre called “bad faith”’. (Ross 1986: 73). This would not be of much significance from the perspective of the *recipient* of testimony, were it not for the fact that, if one accepts the subjective impossibility of offering one’s own words *merely as evidence*, then when we (in our capacity as hearers) do on occasion treat the words of others as mere evidence ‘we are not accepting them in the spirit in which they are offered’ (ibid.). Other authors, by contrast, have resolutely defended an evidential conception of testimonial knowledge. Hence, Stephen Schiffer argues that critics of the evidential view ‘are mistaken about what is required to know something on the basis of an inference to the correct explanation’, since ‘whether knowledge is based on inference from evidence is not about the actual movement of thought, the considerations one actually ponders; it is about the structure of beliefs that sustain one’s conclusions.’ (Schiffer 2001: 2315).

### **Testimonial IBE: the basic idea**

The positions discussed towards the end of the previous section, put forward by Ross (1986) and Schiffer (2001) respectively, provide a good starting point for analysing the prospects of an IBE-based account of testimonial knowledge. One challenge which such an account must address is the kind of sentiment expressed by Ross, according to which an overly evidence-based assessment of testimony is either psychologically implausible or may even occasion a disunity in our testimonial dealings with others, by driving a wedge between what is demanded by the hearer (e.g., conclusive evidence) and what can be reasonably expected from the speaker (e.g., her sincerely giving *assurance* as to the truth of her testimony).<sup>3</sup> The situation is not so different from the one faced by Fricker’s local reductionism, which demanded an inferentialist ‘monitoring’ of the speaker: If monitoring is construed too loosely, one loses out on relevant – and, for any reductionist project, necessary – distinctions, such as between (direct, hence basic) perceptual knowledge and (mediated, hence reducible) testimonial knowledge; however, if monitoring is construed too narrowly, for example by demanding an ongoing conscious effort on the part of the hearer, the proposal becomes psychologically implausible and runs into conflict with our phenomenology of accepting testimony. If one leaves considerations of the phenomenology of belief-formation momentarily aside and takes Schiffer’s response seriously, then the justification of testimony as a source of knowledge need not reside in a careful retracing of any ‘actual movement of thought’, but may be best analysed through a reconstruction of the inferential relationships that sustain the resulting testimony-based beliefs. (Schiffer 2001: 2315). From a perspective

---

<sup>3</sup> For an elaboration of this Assurance View of testimony, see (Moran 2006).

of IBE, the project then becomes one of developing an account of Testimonial Inference to the Best Explanation, or ‘TIBE’, for short (following Lipton 2007: 238).

The basic idea of TIBE can be easily stated: It is the thought that ‘a recipient of testimony (“hearer”) decides whether to believe the claim of the informant (“speaker”) by considering whether the truth of that claim would figure in the best explanation of the fact that the speaker made it.’ (Lipton 2007: 238). By adding only a little more detail to this basic picture, one can posit a sequence of inferential steps, according to which ‘the recipient of testimony is seen as making an inference to the best explanation of why her source – say, John – said that  $p$ : she infers that John said that  $p$  in part because he believes that  $p$ , and she infers that John believes that  $p$  in part because  $p$  is the case.’ (Malmgren 2006: 230). It is this basic step-by-step inferential procedure that has been attributed to, and endorsed by, various philosophers such as Elizabeth Fricker (1994), Jack Lyons (1997), Paul Thagard (2005), and Stephen Schiffer (2001 & 2003). Schiffer gives a nice example of how this more elaborate, step-by-step version of TIBE is supposed to work in the concrete case of a speaker, Sally, informing her interlocutor, Abe, that it is snowing outside:

Thus, when Sally informs Abe that it’s snowing by uttering ‘It’s snowing’, Abe acquires the knowledge that Sally believes that it’s snowing, and this knowledge is crucial to Abe’s coming to know that it’s snowing, and this is in part because part of the explanation of the fact that Sally believes that it’s snowing is that it is snowing.

The rest of the story – the extent to which Abe’s knowledge is inferential – is a matter of debate [...]. (Schiffer 2003: 303)

Why is this example an instance of TIBE? Because, as Schiffer points out, in testimonial encounters of the sort that takes place between Sally and Abe, ‘when we come to know  $p$  on the basis of having been told  $p$ , it is crucial both that we believe that the speaker believes  $p$  and that the best explanation of the speaker’s believing  $p$  essentially includes the fact that  $p$  is true.’ (ibid.). Importantly, on this account, we are under no obligation to gather independent evidence as to the truth or falsity of  $p$ . After all, as Lipton puts it, we want ‘an account of testimonial inference, not testimonial avoidance’ (Lipton 2007: 245). Hence, one aim of TIBE must be to show ‘how the hearer can assess the speaker’s testimony without having independently to determine the truth value of what was said.’ (Lipton 2007: 243). Similarly, we do not accept someone’s testimony that  $p$  *because* we know  $p$  to be true: if we already know that  $p$ , there is no need to take anyone’s word for it. In genuine cases of learning from the word of others, we do not infer from the truth of the matter to the acceptability of the testimony; rather, it is explanatory considerations, together with the fact that the best explanation of the telling *includes* the truth of what we are told, which guide our acceptance of the testimony.

TIBE comes in degrees. It is one thing to demand that the best explanation of an instance of telling entail that what we are told is true, but it is quite another to demand that, when we infer the truth of the testimony, we must always make a prior inference to the belief of the speaker. Yet the latter is what most proponents of IBE-based accounts of testimonial knowledge, including Gilbert Harman in his original reference to testimony as an example of

IBE (see previous section), insist on.<sup>4</sup> Schiffer, for example, calls it a ‘mistake’ to deny that, in cases such as the one quoted above, it must always be ‘necessary that Abe actually believe that Sally had her belief or that she had it because it was snowing’. While Abe ‘need not consciously ponder any thought, [...] he must have the beliefs in question, just as you must actually have the belief that the streets are wet because it rained when you see the streets and instantaneously form the belief that it rained, even though, so to say, that belief need never have consciously crossed your mind.’ (Schiffer 2001: 2315). I am not suggesting that we do not most of the time have tacit, or implicit, beliefs of the sort required by Schiffer, but it seems at least possible to distinguish between a ‘weaker’ version of TIBE, which leaves open the precise way in which we reason from the fact of assertion to the truth of the testimony, and a ‘narrower’ (and hence stronger) version, according to which every instance of TIBE is an instance of making inferences about another’s belief system. I shall return to this question, and why it may be relevant in certain cases of testimonial knowledge, in the next but one section of this paper.

One notable exception to the view that we must always infer the truth of a given testimony via an intermediate inference to the speaker’s belief, is Lipton’s construal of TIBE as simply ‘an abductive inference from the fact of utterance to the fact uttered’ (Lipton 2007: 243). Partly this is because Lipton takes TIBE to be primarily a *descriptive* account of our testimonial practices, which is not so much concerned with whether or not this mode of assessing testimony can reliably tell true from false testimony, but with the question of ‘whether the account illuminates the way we in fact go about deciding whether to believe what we are told’ (ibid.). TIBE, thus understood, is a descriptive framework for our testimonial practices, and as such ‘can be adapted to fit’ different theoretical outlooks, such as a ‘more thoroughgoing inferentialist picture’ (Lipton 2007: 244), without entailing any particular level of restrictiveness regarding our standards of admissibility of testimony. This stands in marked contrast to the proposals of Fricker (1994) and Schiffer (2001), and their endorsements by Malmgren (2006) and others, all of whom aspire to an IBE-style account of testimony that identifies, in a criterial way, the specific mechanisms or inferential relationships that confer ‘knowledge-providing justification’ (Schiffer 2001: 2315) to some testimony-based beliefs rather than others. But Lipton is not entirely alone in pursuing a more descriptive approach to the problem of testimonial IBE; other authors, in a similar vein, have discussed either the way in which credibility features in our explanatory inferences relating to testimony (see Thagard 2005), or the application of explanatorily relevant maxims and countermaxims to testimony concerning a specific domain, such as historical records (see Jardine 2008).

---

<sup>4</sup> For an analogous account of historical testimony, see Jardine’s summary of a view he attributes to Charles Seignobos (1854-1942), according to which our acceptance of historical testimony ‘involves a double reconstruction, of the psychological state of the testifier that is expressed in the testimony, and of the processes which gave rise to that psychological state.’ (Jardine 2008: 165).

### **Interlude: Testimony, IBE, and the evolution of creative thought**

Before considering some of the pros and cons of TIBE in detail in the next section, I briefly want to reflect on a recent suggestion, by Peter Carruthers, concerning the connections between our practice of exchanging testimony, our capacity to make explanatory inferences, and the very possibility of creative thought. Carruthers links the evolution of our inclination towards inferring best explanations to the emergence of language. In making this connection, however, his main concern is not with promoting any particular account of how we accept testimony. Rather, he uses the (purported) evolutionary link between IBE and testimony as an argument in defence of the thesis of *massive modularity* of the human mind. According to this thesis, the human mind consists entirely, or mainly, of mental modules, each of which is adapted to a relatively narrow, domain-specific class of cognitive tasks. (See Carruthers 2006.) One challenge which a theory of mental modularity must meet is to account for certain important and distinctive features of human cognition, of which Carruthers identifies three main categories: ‘*flexibility* of content; *creativity* of content; and *abductive inferences* performed upon such contents.’ (Carruthers 2003: 503). It is the last two of these categories, along with the connections between them and their relation to testimony, that are of interest to us here.

One characteristic feature of human cognitive development is the emergence, early on in our lives, of a capacity ‘to generate, and to reason with, novel suppositions or imaginary scenarios’ (Carruthers 2003: 511), which manifests itself in children as *pretend play*, and carries over into adulthood in the form of *creative thinking*. How could such a capacity – ‘to *suppose* that something is the case (that the banana is a telephone; that the doll is alive), and then think and act within the scope of that supposition’ (ibid.) – have arisen within a modular framework? As Carruthers argues, much of the cognitive activity that accompanies childhood pretend play as well as creative supposition-generation in adults, takes the form of rehearsed ‘inner speech’, through which contents can be globally broadcast across a number of mental modules and, by being placed in novel cognitive contexts, can associatively give rise to new contents. (See Carruthers 2006: 334.) However, we do not merely generate new contents for the sake of novelty, we also do so in order to solve problems and in contexts that require action, and for these ends we must also ‘come to *believe* some of our suppositions’. (Carruthers 2003: 512) Deciding on the best of a number of possible solutions to a problem, or choosing the best of a set of imagined courses of action, engages the very cognitive capacities that are commonly seen to be at work in abductive reasoning. Inferring the most likely outcome of a proposed course of action involves considering a number of hypotheses concerning what will, or might, happen. If it is indeed the case that much of our cognitive life consists in rehearsing various scenarios and hypotheses in ‘inner speech’, then it is perhaps significant, Carruthers argues, that when ‘the hypotheses in question are expressed in language, the problem of inferring the best explanation reduces to the problem of deciding which of the candidate sentences to believe in the circumstances.’ (Carruthers 2006: 364). The problem of deciding between candidate sentences that we are presented with is, of course, a familiar one: it is simply a version of the problem of which testimony to accept, and

when. Hence, according to Carruthers, it seems plausible ‘that the principles of testimony-acceptance are historically and developmentally prior to the principles of inference to the best explanation.’ (Carruthers 2003: 514).

No doubt this move from the problem of IBE to the problem of testimony, through the invocation of the phenomenon of ‘inner speech’, may strike some readers as a little too quick for comfort. But Carruthers does not claim to have given us a conclusive argument. His story is a genealogical one, and as such is meant to shed light on why we have the testimonial and inferential practices we have, and how they might relate to one another. The suggestion is not that we can somehow ‘reverse-engineer’ a justificationist account of IBE in terms of testimony, or vice versa. Rather, what this speculative genealogical story of the emergence of IBE shows is that there is room for an account of testimonial acceptance that takes the reality of our inferential practices seriously without reducing testimony to just another application of an overly rigid inferentialist framework. According to the picture presented by Carruthers, testimony comes first, IBE comes second, and the latter takes its cues from the former.

Importantly, this does not mean that testimonial acceptance is indiscriminate. On the contrary: if testimony (or, to be precise: some genealogically prior practice of ‘proto-testimony’) is supposed to have functioned as a model for IBE, then it must have been selective in discriminating between acceptable and unacceptable testimony, in much the same way that full-fledged IBE is meant to discriminate between bad, better, and best explanations. Which factors, then, could have been responsible for the relative success and reliability of early testimonial practices that made them a good model for more ‘advanced’ IBE-style inferential patterns? Two kinds of factors can be ruled out: innate biases (such as Thomas Reid’s matching ‘Principles of Credulity/Veracity’), and ‘first-order’ inferential mechanisms of the sort proposed by local reductionism (which are thought to ‘monitor’ each individual speaker for natural signs of insincerity or incompetence). Innate biases may well be present, and even ‘*causally* necessary’ for our acquisition of testimony-based beliefs, but, as Lyons points out, ‘this causal responsibility does not imply *justificational* responsibility’ (Lyons 1997: 165). In addition, the presumed innateness of such biases seems generally orthogonal to the emergence of abductive modes of justification. (On this point see also Lyons 1997.) ‘First-order’ monitoring mechanisms of the kind posited by Fricker’s local reductionism also fail in this regard. What makes such mechanisms ‘first-order’ is their presumed ability to identify, through a process of filtering out possible defeaters, a *specific property of the speaker* – the ‘weakest gap-bridging property’ (Fricker 1994: 129) – that bridges the gap between *mere say-so* of the speaker and *truthful belief* formed on the basis of the speaker’s testimony. Contrary to what local reductionists have claimed, such knowledge is highly specific and not easy to come by. (On this point, see Gelfert 2009: 176-181.) Even if reliable mechanisms of this sort could plausibly exist – which is far from clear – they would, in some sense, be achieving *too much*: their brute reliability would render any comparative assessment of the acceptability of a given testimony unnecessary, yet it was such comparative judgments that were thought to serve as a model for IBE.

Luckily, in our assessments of testimony we can help ourselves to an alternative basis of evidential considerations. This basis has the advantage that it is easier to come by than the

highly specific first-order evidence demanded by local reductionism, while not requiring the gratuitous postulation of new innate faculties. It derives from considerations of *relevance*, where the term is here understood in the technical sense due to relevance theory, as developed by Dan Sperber and Deirdre Wilson (1995). A central ingredient in Sperber and Wilson's account of discourse interpretation is the mutual presumption of relevance: both interlocutors are committed to making their interaction as economical and informative as possible. This applies at the level of the speaker, where it corresponds to some aspects of the Gricean picture of speaker-intentions: In being ostensibly addressed to an audience, the speaker's utterance automatically conveys the presumption of its own optimal relevance. But it also applies at the level of the hearer, who applies the same twin principles of economy and informativeness – summed up in the maxims 'minimize processing effort' and 'maximize information gained' – to the interpretation of the speaker's utterance. Note that the desiderata expressed by these principles, or maxims, concern facts about the hearer's *own* cognitive processes: how easy it is for him to extract information from a given testimony, and how much information he is able to extract. While this means that considerations of relevance depend on the hearer's background beliefs and other contextual factors, it also means that evidence of relevance is readily available: the hearer need only monitor his *own* cognitive goings-on. Our interpretations and deliberations about what other people say do come to an end, and typically they do so when we have subjectively exhausted their informational potential. There is no guarantee that a hearer will always end up with the objectively most accurate or informative interpretation; rather, in our role as hearers, we tend to adopt a *satisficing* policy, governed by heuristics and other cognitive shortcuts. (See also Carruthers 2006: 370.) Important examples of such cognitive strategies would be checking testimonial statements (i.e., potential testimony-based beliefs) for coherence with our background beliefs, or comparing new instances of testimony with past (remembered) instances of testimony: if the reports of people (on the same topic, or in general) cohere with each other and with our own observations and background beliefs, this provides considerable *abductive* grounds for accepting what we are told. Significantly, the justification that derives from such coherence is itself *non-testimonial* in character: *that* testimonies cohere with one another, and do not contradict our background beliefs, is something that, if true, we can ascertain firsthand and without reliance on direct 'first order' evidence regarding properties of the speaker. Typically, this happens as a matter of course when we register and understand utterances as acts of communication. It thus seems plausible to co-opt the principle of relevance as a principle of testimony-acceptance. Hence, at least 'in default circumstances or circumstances where the credibility of an informant isn't in question', we do well to adopt the following maxim: '*believe the interpretation of the other's utterance that is simplest and most informative*' (Carruthers 2003: 516).

### **TIBE: problems and attractions**

As a form of Inference to the Best Explanation, TIBE helps itself to a powerful general account of non-demonstrative, or inductive, inference. In doing so, however, it also inherits any problems and open questions that may be associated with IBE in general. In the present section, the focus is mainly on problems and attractions of *testimonial* IBE, rather than IBE *in general*. Given that TIBE is a theoretical framework for analysing testimonial knowledge, rather than a single philosophical position, and as such allows for multiple specific formulations, part of the discussion will necessarily consist in spelling out possible versions of TIBE. For, as Nick Jardine puts it, the basic criterion advocated by TIBE – that ‘the production of the testimony is best explained by an account which makes probable its truth’ – stands in need of some elaboration. (Jardine 2008: 167).

Turning first to the attractions of TIBE as an account of testimonial knowledge, perhaps the most important positive feature is the fact that it aims to explain our testimonial practices in a way that is both descriptively adequate and does not posit unnecessary new mechanisms. In particular, it posits neither testimony-specific ‘innate biases’ (as Reid did) nor any new species of epistemic justification (as Burge does when he argues for a weak kind of ‘epistemic entitlement’ that is largely specific to testimony and memory as sources of knowledge). Instead, TIBE intends to show ‘that warranted testimonial beliefs are based on rules of inference or mechanisms of belief acquisition that apply to the beliefs from various sources, not just the source of testimony.’ (Lipton 1998: 24). As a methodological corollary, this implies that, in our attempts to construct a philosophical account of testimonial knowledge, we ought first to explore how far we can get, *knowing what we do* about our cognitive apparatus, before positing any qualitatively new cognitive mechanisms or principles. Interestingly, some accounts that claim an affinity to TIBE, appear to violate this demand, most notably Fricker’s local reductionism which posits subconscious monitoring mechanisms that would furnish the kind of first-order ‘knowledge-providing’ evidence that, as I argued in the previous section, can be seen to be excessive and perhaps even self-defeating. (For an extended argument to this effect, see Gelfert 2009.) At the same time that TIBE aspires to be economical concerning the kinds of cognitive mechanisms and principles of justification it requires, it also aims to be descriptively adequate. In particular, to use Lipton’s earlier phrase, it aims to give ‘an account of testimonial inference, not testimonial avoidance’ (Lipton 2007: 245). This sets TIBE apart from revisionist accounts of testimony, such as Hume’s global reductionism, according to which we ought to rely on testimony only to the extent that we have regularly observed conformity between testimony and the facts. Whereas revisionist accounts of testimony attempt to *correct* our testimonial practices (usually by imposing limitations), TIBE takes the overall legitimacy of our reliance on testimony for granted and attempts to *vindicate* it.

While the TIBE approach is attractive for the reasons just discussed, it nonetheless has its fair share of problems. In the remainder of this section, the focus will once again be on issues specific to testimonial IBE, rather than on problems with inference to the best explanation in general. In particular, I want to move from criticisms of the ‘narrow’ version

of TIBE (in the sense discussed in the section preceding the previous one) to challenges and proposed amendments of ‘weaker’ versions of TIBE. Recall that, according to the ‘narrow’ version, every instance of TIBE is an instance of making inferences about the speaker’s belief system: ‘The best explanation for why the informant asserts that  $p$  is normally that, first and most relevant here, he believes it for duly responsible reasons and, second, he intends that I shall believe it too (by virtue of recovering this intention from his assertion).’ (Adler 1994: 274) The ‘weaker’ conception of TIBE is only committed to the basic criterion mentioned at the beginning of this section – that the production of the testimony that  $p$  is best explained by an account that makes it probable that  $p$  – and not to any specific picture of the speaker’s intentions and beliefs.

As a first challenge, I wish to offer a novel criticism of ‘narrow’ TIBE on the account that, in certain cases that it should be able to accommodate, it does not allow for testimony to be a *generative* source of knowledge. As Jennifer Lackey (1999) has shown, one can easily construct cases (however rare and otherwise unusual) where the hearer comes to know  $p$  by accepting what he has been told, even though the speaker failed to know that  $p$ , because of the presence of a defeater in her belief system. As Lackey argues, testimony can fail to transmit doxastic defeaters, while successfully transmitting the speaker’s relevant true belief in such a way as to bring about knowledge in the hearer (provided the hearer does not independently suffer from the same, or any other, defeaters to the corresponding belief). Consider Lackey’s example of Mrs. Smith, a conscientious biology teacher who is also a devout creationist. She has thoroughly researched the theory of evolution, using a range of reliable scientific sources and other available evidence, yet none of this has managed to convince her that evolution is more than ‘just a theory’, let alone a correct one. However, she dutifully adheres to the curriculum and teaches her students about the scientific facts of evolution, keeping her reservations to herself. In doing so, she arguably imparts knowledge to the students in her class, who are blissfully unaware of their teacher’s being a closet creationist.

Lackey’s scenario, in which the non-transmission of doxastic defeaters allows us to credit the hearer, but not the speaker, with knowledge, is intended as an argument against a purely *transmissive* model of testimonial knowledge. However, it can also be applied, with only minor modifications, to the ‘narrow’ (belief-centred) model of TIBE. One might argue that, as the case has been described, the teacher’s secret belief in creationism does not preclude that it is a clear instance of TIBE: *from the students’ point of view*, the teacher’s classroom testimony may still be best explained by the assumption that she believes what she teaches and that she intends that her students believe it too. Hence, the inference from the act of telling to the truth of the testimony, via the attribution of corresponding beliefs and intentions to the speaker, might seem immune against hidden doxastic defeaters. Consider, however, a slight modification of the example. Imagine that, unlike in the previous case, the students in Mrs. Smith’s class know about her private endorsement of creationism. (Perhaps someone spotted her signature on a petition demanding that creationism be taught in schools, or overheard her saying to a friend how much she dislikes having to teach her students a ‘false’ scientific theory.) Let us further assume they also know first-hand that, in the past,

their teacher never let her personal opinion influence her teaching, and they have ample evidence of the teacher's professionalism. Clearly, in the modified version, the teacher's doxastic defeaters are no longer hidden from view, so her students would be unjustified in inferring that she actually believes the scientific facts she reports in her teaching. But this need not mean that the students cannot rely on the teacher's testimony: given the classroom setting and the constraints that govern testimony in this context, and given that the students have independent reasons not to doubt the teacher's professionalism, they would be well-advised to accept what she teaches, thereby acquiring knowledge of the scientific facts. It appears, then, that 'narrow' TIBE, which proceeds via inference to the belief of the speaker, would indeed be *too narrow*: it would rule out perfectly legitimate ways of acquiring knowledge from a reliable testifier. Note, though, that this objection does not apply to TIBE *in general*. In particular, it does not count against the 'weaker' version of TIBE, which is only committed to the basic criterion that an act of telling is best explained by an account that makes probable the truth of what is told. In the example of Mrs. Smith, both in its original and its modified form, the best explanation of her classroom testimony might appeal to her professional role and track record as a teacher, the existence of curricula that govern what is taught in schools, and social mechanisms that ensure that someone who misused her role as a teacher to indoctrinate children would not remain employed for very long. It is such background knowledge that enables us to tell a vindictory story of why the specific testimony is indeed acceptable.

What mattered in the closet creationist example, was the peculiar role of (initially hidden) doxastic defeaters and, in the modified version, the observation that the context in which a claim is made (in this case: the institutional setting with its constraints on classroom testimony) can overrule defeaters even when they are 'out in the open'. What is a defeater for the speaker, need not be a defeater for the hearer. The fact that the example involved a *scientific* claim was only of secondary importance. But scientific testimony arguably comes with baggage of its own and, I want to suggest, poses an independent challenge to IBE-based accounts of testimonial knowledge, even when these are understood in the 'weaker' sense discussed above. Consider any fundamental statement of science that is communicated via testimony, e.g. the law of gravity. What would it mean to say, as the basic criterion of TIBE would have it, that the instance of testimony in question is best explained by an account that makes the truth of what is told – in this case, the statement of the law of gravity – probable? Whereas in cases of ordinary empirical testimony (such as eyewitness reports), we can easily conceive of many causal stories that would link the circumstances of its production (including the mere fact that it has been produced) to the truth of the testimony, these connections seem a lot less direct and more elusive in the case of testimony that asserts general, universal, or abstract truths.<sup>5</sup> If 'narrow' TIBE is too narrow because it links testimonial knowledge too closely to what can, or cannot, be inferred about the speaker's beliefs, then perhaps 'broad'

---

<sup>5</sup> Abstract (e.g., mathematical) truths pose special problems in relation to testimony, but the more general point also applies, for example, to distant historical testimonies: How far back in time do our inferred explanations, of testimony regarding distant historical events, have to reach, in order for them to qualify as 'best', or at least 'good enough'?

TIBE runs the risk of construing the link, between the fact of assertion and the truth of what is being asserted, too loosely. We may simply be unable to tell whether the asserted fact (e.g., that the law of gravity obtains) is part of the *best explanation* of why someone asserted it.

The challenge, then, is for TIBE to identify a ‘middle ground’ of considerations that are explanatorily relevant to the testimony in question, but that neither make acceptability of the testimony dependent on speculations about the speaker’s psychology, nor make excessive demands concerning our understanding of, and our ability to assess, the various indirect ways in which the asserted fact (provided what has been asserted is indeed true) could manifest itself. Luckily, it is highly plausible that, under normal circumstances, we do have access to such considerations. In fact, one class of considerations consists precisely in the kind of background knowledge of social institutions that we encountered in the case of the creationist biology teacher. The kind of testimony that students receive from their teachers is best explained by the fact that schools are designed to be *institutions of learning*, where knowledge that has been carefully vetted and has been formally enshrined in curricula is passed on to students. The situation is slightly different in the case of scientific testimony, where one often deals with novel claims rather than established knowledge. While we are often in no position to assess how the truth of a particular scientific claim would specifically feature in the best explanation of the production of the corresponding scientific testimony, we do know a lot about the *kinds* of processes that are responsible for our encountering such scientific testimony. The former would require specialist scientific knowledge (and detailed knowledge of the social organisation of science), whereas the latter merely requires a certain degree of familiarity with how science operates and how it communicates its results (e.g., through peer review and publication in highly competitive journals); the latter is easy to obtain, the former is not.

As an example, imagine I read a news report about the 2008 Nobel Prize in Medicine, which includes the statement that part of the prize is awarded for the discovery that the human papilloma virus is the primary cause of cervical cancer. Presumably, at some level, it may be possible *in principle* to construct an explanation in which the truth of the claim in question – the hypothesis that cervical cancer is due to HPV – is directly responsible for the production of the news report I encounter. But we do not normally reason like this, and we do not need to: all that is required is that, given what we know about how Nobel prizes are usually awarded, we can infer that if there wasn’t overwhelming evidence for the hypothesis, it would not have merited the Nobel prize and we would not have encountered the news report in question. Perhaps this sounds obvious or even trivial, but it does tie in nicely with an observation mentioned at the end of the previous section: namely, that some of the most powerful abductive evidence for the reliability of testimony consists in the coherence of individual testimonies with one another, as well as with our observations and background beliefs, including our background knowledge about the processes underlying the production of said testimonies. As long as we can reasonably presuppose that, in each successful case of accepting testimony, it is possible *in principle* to give an explanatory account of the act of assertion that makes probable the truth of what is asserted, we can discharge this

responsibility by making appropriate inferences based on the context in which the testimony was made.

### **The place of IBE in assessing testimony: an irenic proposal**

As mentioned earlier, TIBE is explicitly intended to be descriptively adequate with respect to our actual testimonial practices – more so, perhaps, than any other position within the epistemology of testimony. But, of course, any account of testimonial knowledge that was utterly disconnected from the actual ways in which we adjudicate, and respond to, testimony would be faced with some tough questions. At the very least, a plausible account should attempt to accommodate the fact that our testimonial encounters include both instances of acceptance and instances of rejection. And indeed, this is precisely what one finds in the philosophical literature. In one of the early controversies within the contemporary debate about testimony, Tony Coady and Elizabeth Fricker exchanged criticism over an example discussed by Coady, which invites us to reflect on how we react when we call, say, the telephone company to find out about an invoice we forgot to pay: ‘[I] am told by an anonymous voice that it comes to AU\$165 and is due on 15 June.’ For Coady, who is a non-reductionist about testimonial justification, the situation is clear: ‘No thought of determining the veracity and reliability of the witness occurs to me nor, given that the total is within tolerable limits, does the balancing of probabilities figure in my acceptance.’ (Coady 1992: 143) Fricker, by contrast, points to the clause ‘...given that the total is within tolerable limits...’ as evidence that there must be ‘active sub-personal monitoring of the speaker by the hearer for signs of lack of sincerity or competence’ (Fricker 1995: 405), since surely we would not accept just any figure the voice tells us. Whatever one’s fundamental intuitions regarding testimony, whether one is a reductionist or anti-reductionist, it is obvious that a good account of testimony must be able to account for paradigmatic cases of simple unquestioned acceptance as well as for obviously legitimate cases of rejecting implausible testimony. In the remainder of this paper, I want to develop an account that can reconcile the two opposing attitudes in the above example in an irenic spirit, by subsuming both under a TIBE-based framework.

A first important realization is that, while the attitudes exemplified by Coady and Fricker are indeed antagonistic, at the same time they correspond to two very real desiderata, which may be called *trusted acceptance* and *rational rejection*. (See Gelfert 2008: 41-43.) The simple idea here is that our predicament with respect to testimony is such that it is often rational to trust what we are told, and that an attitude of simple acceptance is justified in such circumstances, but that it must also, on occasion, be rational to reject testimony. Without either of these two desiderata, an account of testimony would necessarily have to be regarded as incomplete. Both trusted acceptance and rational rejection have been variously endorsed by prominent contributors to the epistemology of testimony, although the usual move has been to stress one at the expense of the other. Thus, trusted acceptance as a theoretical desideratum has been emphasised by John McDowell who argues that at ‘the core of a good

general account of testimony’ should be the intuition that, ‘if a knowledgeable speaker gives intelligible expression to his knowledge, it may become available at second hand to those who understand what he says.’ (McDowell 1998: 417). By contrast, Fricker clearly points to the desideratum of rational rejection when she writes that ‘we know too well how, and how easily, what we are told may fail to be true’ and cautions us not to ‘trust anyone, let alone everyone, uncritically’ (Fricker 1995: 400). Instead of regarding trusted acceptance and rational rejection as defeasible desiderata of an account of testimony, however, reductionists and anti-reductionists attempt to assimilate our reaction to testimony to one or the other, thereby getting locked in a dispute over which should take priority.

TIBE fares better on this score, partly because TIBE-based accounts have a natural affinity, and often explicitly endorse, what Lipton has called ‘a “default-trigger” model of testimony’ (Lipton 2007: 240). On the simplest version of such a model, while ‘in most contexts the hearer simply accepts what she is told, without engaging in any conscious evaluation or inference’ (ibid.), there are also circumstances that may trigger a ‘switch from default into evaluative mode’ (Lipton 2007: 250), where the hearer considers whether she should believe the speaker. This is similar to the model developed by Paul Thagard, who postulates both a ‘default pathway’ in which, all else being equal, hearers ‘more or less automatically respond to a claim by accepting it’, and a ‘reflective pathway’ in which hearers ‘evaluate the claim based on its explanatory coherence with everything else they believe’ (Thagard 2005: 295). Thagard distinguishes, at a descriptive level, four specific kinds of ‘reflection triggers’: ‘lack of credibility of the source, non-credible behavior of the source, inconsistency of the claim with other beliefs, and incompatibility of the claim with the hearer’s goals’, all of which – with the possible exception of the last – also provide, prescriptively, ‘good grounds for additional, more reflective processing’ (Thagard 2005: 298). Finally, consider Jardine’s recent discussion of maxim-based accounts of testimony, for which there is a long tradition throughout the history of logic and rhetoric. What is striking about such historical accounts is that one finds a ‘pairing of maxims for the acceptance of testimony with countermaxims for its rejection’, and, as Jardine notes, ‘a theory of testimony owes us some indication of why such contrary maxims seem plausible to us’ (Jardine 2008: 163). All three authors – Lipton, Thagard, and Jardine – are proponents of (different versions of) TIBE, and are concerned more with the *management* of testimony than with the question of whether or not testimonial justification is reducible to other epistemic sources. In the light of our earlier discussion, one can then attempt an answer to the challenge formulated by Jardine: the simultaneous existence of maxims and countermaxims for the management of testimony is unsurprising, precisely because trusted acceptance and rational rejection are both desiderata of any plausible theory of testimony.

And yet, there remains something of a discrepancy between TIBE’s acknowledgment of our thoroughgoing reliance on testimony and the account it gives of how we make specific inferences to the truth of a testimony. We have already seen how ‘narrow’ TIBE fails to account for certain instances of testimonial knowledge, as in the case of the non-transmission of doxastic defeaters. Also, I argued that the basic criterion of ‘weak’ TIBE – that an act of telling must be best explained by an account that makes probable the truth of what we are told

– may be too unspecific: it posits a line of inferential reasoning – from the telling to the truth of what is told – that may be available to us *in principle*, but whose merits we often lack the expertise to assess. But if we lack the ability to *actually* formulate and assess those explanations that are relevant to the content of the testimony in question, then the idea that the *best* among those explanations would entail the truth of the testimony becomes little more than an article of faith. TIBE would still owe us an explanation of why we are justified in following a ‘default pathway’ of simple acceptance in most contexts.

In order to remedy this situation, I wish to propose a modified role for IBE in the assessment of testimony, one that resolves the discrepancy described in the previous paragraph, while at the same time staying true to the goal of accounting for our *actual* testimonial practices – which include both trusted acceptance and rational rejection. To this end, I want to propose a *dual role* for IBE in how we adjudicate testimony. In particular, I want to suggest that IBE has a specific role to play in the *rejection* of a given testimony: it is when the best explanation of the fact of assertion makes what is asserted improbable, or unreliable, that we should reject it.<sup>6</sup> At first sight, this might appear implausible, given that it looks like a mere inversion of standard accounts of TIBE, according to which we should base *acceptance* of a given testimony on the fact that the best explanation of its production entails, or makes probable, its truth. However, from the point of view of an epistemic agent, both the rejection of a (potential) false belief and the addition of a (new) true belief are epistemic goals. Hence, both the standard version of TIBE and the revised version I am proposing – according to which we employ IBE in order to identify specific testimonies that are best rejected – are in the service of improving the epistemic position of the recipient of testimony. When does the best explanation of an assertion render what is asserted improbable or unreliable? A good example would be unsolicited testimony that promises a (deferred) profit in exchange for a (present) cost.<sup>7</sup> Consider the following example, due to Lipton:

A man rang my doorbell and claimed that my rain gutters are loose. Should I believe him? They look fine to me, I know that he hasn’t been up on the roof to inspect them properly, and I am further discouraged by the fact that he wants me to pay him today to fix them tomorrow. (Lipton 2007: 244)

Clearly, in this case, it is rational to reject the testimony in question, and we are justified in rejecting it precisely because the best explanation of the testimony entails that it is probably false. The point is this: the very factors that, in Thagard’s terminology, function as ‘reflection triggers’, at the same time often provide abductive evidence for the falsity, or unreliability, of the testimony they accompany. (One obvious example would be the unsolicited denial, from a child, of a misdeed that has not yet been found out: the child’s ‘I didn’t do it’ may be the best initial evidence that something is wrong.) Also, if it is indeed the case, as seems psychologically plausible, that the ‘reflective’ or ‘evaluative’ mode is something that *needs to be triggered*, then one would expect there to be specific inferential criteria in place for such

---

<sup>6</sup> Precisely how reliability, or the lack thereof, is to be construed – whether as a brute statistical fact, or in modal terms, as the sensitivity or insensitivity of the belief in question – may itself depend on the context and the epistemic standards appropriate to it.

<sup>7</sup> Nothing much hinges on the fact that the example involves a speech act of promising, since it is the constative testimony in the run-up to a promise that is important here.

triggered cases (rather than a mere extension of the criteria that underlie cases of unreflective acceptance). In other words, we typically reject testimony for reasons that are different in kind from those that underlie our overall tendency to accept what others tell us. This is precisely what the modified version of TIBE acknowledges when it demands that, in the ‘reflective’ or ‘evaluative’ mode, we screen testimonies for likely deficiencies – either in the sense that they are likely to be false, or because they may be unreliably produced.

As emphasised throughout this section, rejection of testimony takes place against the backdrop of our pervasive reliance on testimony, for which trusted acceptance is indispensable. Whereas, on the model I am proposing, instances of *rejection* occur as the result of a specific act of IBE – namely, when the best explanation of the telling makes it *improbable* that what we are told is reliable – our *acceptance* of testimony, I want to suggest, is based on the general reasoning discussed at the end of previous section: i.e., on the observation that some of the most powerful abductive evidence for the overall reliability of testimony consists in the coherence among the various sources we draw on, as well as on the coherence of new testimonies with our observations and background beliefs. IBE thus has a dual role to play: on the one hand, considerations of coherence, along with the success of our various testimony-dependent collective projects (such as education, science, or history), abductively sustain a stance of *default acceptance*, whereas, on the other hand, we are justified in *rejecting* a given testimony whenever the best explanation of the circumstances of its production (including the fact that it has been produced) casts doubt on the reliability of what we are told. Hence, when construed properly, in a way that acknowledges its dual role in the management of testimony, TIBE can help us see why we can often take on trust what others tell us, without thereby being gullible.

## References

- Adler, J.E. (1994). Testimony, trust, knowing. *The Journal of Philosophy*, 91, 264-275.
- Adler, J.E. (2008). Epistemological problems of testimony. In E.N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy (Fall 2008 Edition)*.  
<http://plato.stanford.edu/archives/fall2008/entries/testimony-episprob/>. (Accessed 1 November 2008)
- Burge, T. (1993). Content preservation. *The Philosophical Review*, 102, 457-488.
- Carruthers, P. (2003). On Fodor’s problem. *Mind & Language*, 18, 502-523.
- Carruthers, P. (2006). *The Architecture of the Mind: Massive Modularity and the Flexibility of Thought*. Oxford: Oxford University Press.
- Coady, C.A.J. (1992). *Testimony: A Philosophical Study*. Oxford: Oxford University Press.
- Fricker, E. (1994). Against gullibility. In B.K. Matilal, & A. Chakrabarti (Eds.), *Knowing From Words* (pp. 125-162). Dordrecht: Kluwer.
- Fricker, E. (1995). Telling and trusting: Reductionism and anti-reductionism in the epistemology of testimony. *Mind*, 104, 393-411.

- Gelfert, A. (2006). Kant on testimony. *British Journal for the History of Philosophy*, 14, 627-652.
- Gelfert, A. (2008). Learning from testimony: Cognitive cultures and the epistemic status of testimony-based beliefs. In T. Botz-Bornstein (Ed.), *Culture, Nature, Memes: Dynamic Cognitive Theories* (pp. 34-56). Newcastle: CSP.
- Gelfert, A. (2009). Indefensible middle ground for local reductionism about testimony. *Ratio*, forthcoming.
- Grice, H.P. (1957). Meaning. *The Philosophical Review*, 66, 377-388.
- Grice, H.P. (1969). Utterer's meaning and intentions. *The Philosophical Review*, 78, 147-177.
- Harman, G. (1965). The inference to the best explanation. *The Philosophical Review*, 74, 88-95.
- Hume, D. (1999). *An Enquiry Concerning Human Understanding* (T. L. Beauchamp, Ed.). Oxford: Oxford University Press. (First published 1748.)
- Jardine, N. (2008). Explanatory genealogies and historical testimony. *Episteme*, 5, 160-179.
- Lackey, J. (1999). Testimonial knowledge and transmission. *The Philosophical Quarterly*, 49, 471-490.
- Lipton, P. (1st ed. 1991/2nd ed. 2004). *Inference to the Best Explanation*. London: Routledge.
- Lipton, P. (1998). The epistemology of testimony. *Studies in History and Philosophy of Science*, 29, 1-31.
- Lipton, P. (2007). Alien abduction: Inference to the best explanation and the management of testimony. *Episteme*, 4, 238-251.
- Lyons, J. (1997). Testimony, induction, and folk psychology. *Australasian Journal of Philosophy*, 75, 163-178.
- McDowell, J. (1998). Knowledge from hearsay. In J. McDowell, *Meaning, Knowledge & Reality* (pp. 414-444). Cambridge MA: Harvard University Press.
- Malmgren, A.-S. (2006). Is there a priori knowledge by testimony? *Philosophical Review*, 115, 199-241.
- Moran, R. (2006). Getting told and being believed. In J. Lackey, & E. Sosa (Eds.), *The Epistemology of Testimony* (pp. 272-306). Oxford: Oxford University Press.
- Paavola, S. (2006). Hansonian and Harmanian abduction as models of discovery. *International Studies in the Philosophy of Science*, 20, 93-108.
- Reid, T. (1997). *An Inquiry into the Human Mind* (D. R. Brooks, Ed.). Edinburgh: Edinburgh University Press. (First published 1764.)
- Ross, A. (1986). Why do we believe what we are told? *Ratio*, 28, 69-88.
- Schiffer, S. (2001). Communication: Philosophical aspects. In N.J. Smelser, & P.B. Baltes (Eds.), *The International Encyclopedia of the Social Sciences* (pp. 2311-2316). Amsterdam: Elsevier.
- Schiffer, S. (2003). *The Things We Mean*. Oxford: Oxford University Press.
- Sperber, D., & Wilson, D. (1995). *Relevance: Communication and Cognition* (2nd edn.). Oxford: Blackwell.
- Thagard, P. (2005). Testimony, credibility, and explanatory coherence. *Erkenntnis*, 63, 295-316.