Against Credibility

Joseph Shieber

Abstract

How does the monitoring of a testifier’s credibility by recipients of testimony bear upon the epistemic license accruing to a recipient’s belief in the testifier’s communications?1 According to an intuitive and philosophically influential conception, licensed acceptance of testimony requires that recipients of testimony monitor testifiers with respect to their credibility. I argue that this conception, however, proves to be untenable when confronted with the wealth of empirical evidence bearing on the ways in which testifiers and their interlocutors actually interact.

Keywords: Testimony, Monitoring, Knowledge, Trustworthiness, Deception

1. Introduction

Let us consider the following thesis to constitute a strong anti-reductionist position with respect to the epistemology of testimony: if a subject S (seemingly) comprehends a (seeming) testimony by a (seeming) testifier that \( p \), and if that testimony causes or sustains in the normal way S’s belief that \( p \), then S’s belief that \( p \) is thereby prima facie justified. [Cf. Graham 2006: 108] Though strong anti-reductionism found some early and influential adherents, particularly in the apriorist form advocated perhaps most

---

1 I presented an earlier, shorter version of this paper at a symposium at the 2010 Central Division APA meetings, chaired by Edward Hinchman and with comments on the paper by Peter Graham and Elizabeth Fricker. I owe thanks to them, as well as to Alvin Goldman, Hilary Kornblith, Sanford Goldberg, and Ernest Sosa, all of whom gave me feedback on earlier versions of this paper. Two anonymous referees at this journal gave me many helpful suggestions, resulting in a much stronger version of the paper; thanks are due to them as well.
significantly by Burge [1993], most recent theorists have rejected it as too permissive. Indeed, to borrow the title of Elizabeth Fricker’s [1994], it would not be too much of an exaggeration to see the response to anti-reductionism as a reaction ‘against gullibility.’

In reacting against permissiveness, more recent theorists have tended to emphasize requirements intended to impose upon recipients of testimony a responsibility to avoid gullibility. Thus, e.g., it is common in the contemporary discussion of testimonial warrant to suggest that recipients of testimony must be vigilant with respect to a testifier’s credibility. So far, however, the contemporary philosophical discussion concerning the necessary conditions for the licensed acceptance of testimony has been dominated by armchair reasoning concerning those conditions. This is particularly troubling, given the wealth of social psychological research concerning the ways in which individuals actually form beliefs on the basis of their interactions with others.

To see that this lack of awareness of current social psychological research ought to be troubling for those interested in achieving progress in the epistemology of testimony, note that the following three propositions form an inconsistent triad:

(1.1) The overwhelming majority of subjects are not reliable in monitoring their interlocutors for trustworthiness, deceit, and competence.

(1.2) A necessary condition for acquiring warrant from testimony is that subjects reliably monitor their interlocutors for trustworthiness, deceit, or competence.

(1.3) The overwhelming majority of subjects do, at least occasionally, acquire warrant from testimony.

As even a brief survey of contemporary positions in the epistemology of testimony will demonstrate, however, almost all such positions place some onus on recipients of testimony to monitor the credibility of their interlocutors—that is, almost all contemporary positions embrace thesis (1.2). Given that even a cursory reading of the social psychological literature that bears on the ways in which subjects form beliefs on the basis of testimony would seem to provide unimpeachably strong evidence in favour of thesis (1.1), however, it would seem that such contemporary positions retain an allegiance to thesis (1.2) only at the cost of abandoning thesis (1.3) and embracing a very robust scepticism with respect to the prevalence of testimonially acquired warrant.

We will proceed as follows. In section 2 we sketch the ways in which a number of reactions to strong anti-reductionism in the epistemology of

\footnote{For an argument that Kant was an early apriorist, anti-reductionist precursor of Burge, see Shieber [2010a].}
testimony—whether from reductionists, modified anti-reductionists, or from ‘hybrid’ theorists—place demands on recipients of testimony to monitor the credibility of testifiers, either through a requirement that recipients form positive judgments that their interlocutors are trustworthy or competent or through a more minimal requirement that recipients at least be sensitive to evidence that their interlocutors are deceptive. Then, in sections 3 and 4, we survey the psychological literature regarding the prospects for assessing trustworthiness or detecting deception, suggesting that the literature offers little hope that either form of monitoring is reliable. Finally, in sections 5 – 7, we discuss the challenges posed by the empirical data for even modified forms of the theories of testimony surveyed here, paying particular attention to the prospects for various forms of reductionism with respect to testimonial warrant.

2. Evidence for Philosophers’ Acceptance of Monitoring as a Necessary Condition for Testimonial Warrant

In perhaps the earliest contemporary explication of a reductionist rejoinder to anti-reductionist accounts of testimonial justification, Elizabeth Fricker suggests that, in order for a subject’s belief on the basis of testimony to be justified, ‘the subject’s belief must be causally dependent on his belief in (or some kind of sub-doxastic registering of) propositions which in fact constitute the premises of a justification of his belief—which, if articulated, would constitute a good argument for it,’ [E. Fricker 1987: 61] and, in particular, that ‘a valid argument from “He asserted that P” to “P” needs as further premises the ancillary information that the speaker was, on the occasion in question, both sincere, and competent about its subject matter.’ [E. Fricker 1987: 73] Fricker refers to the idea that licensed acceptance of testimony requires belief in—or subdoxastic rendering of—the sincerity and competence of speakers as a ‘monitoring requirement,’ and considers the adoption of such a requirement necessary if theorists concerning the epistemology of testimony are to avoid licensing gullibility. In a recent defence of this position, Fricker [2006] has returned to this claim, underscoring that ‘a proposed epistemology of testimony which posits a presumptive right to trust with no monitoring requirement will license a bad form of gullibility, thereby failing an adequacy condition on an epistemology of testimony.’ [E. Fricker 2006: 618]5

5 Miranda Fricker [1998] provides a further example in her discussion of Edward Craig’s [1990] suggestion that good informants are characterized by their possession of certain indicator-properties. There, M. Fricker notes that ‘surely, what the inquirer needs is someone he can pick out as likely to tell him truly whether p—someone likely to be able and willing to give him the information he wants. So the method of investigating knowledge from the inquirer’s perspective should (or, at least, can) be taken to require that not only the informant’s competence be recognizable, but also her..."
Crucially, many theories that subscribe to reductionist conditions on testimonial justification—even if those theories are not, themselves, reductionist—include necessary conditions involving judgments of credibility. Thus, Paul Faulkner [2000] has written in support of a hybrid theory of testimonial justification, according to which ‘a judgment of credibility … is not irrelevant to the warrant of the associated testimonial belief.’ [Faulkner 2000: 594] Faulkner’s discussion is a very subtle one, but he seems to be suggesting that some minimal judgment of the credibility of the testifier is required in order for the recipient of the testimony to be warranted in believing on the basis of that testimony—this, presumably, is precisely the point in Faulkner’s dubbing his view to be a hybrid between reductionism and antireductionism.4

An analogous account may be found in Lackey’s [2006a] version of a hybrid theory, dubbed ‘dualism,’ which, according to Lackey, claims ‘that testimonial justification requires positive epistemic contributions from both the speaker and the hearer.’ [Lackey 2006a: 171] The positive epistemic contributions from hearers that Lackey foresees involve a requirement to adduce positive reasons for accepting instances of testimony. Again, for Lackey, as for the Frickers and Faulkner, chief among these positive reasons are data involving sincerity and competence of testifiers. Thus, Lackey suggests that one will likely have ‘a substantial amount of inductive evidence for believing that … reports made with sustained eye contact are typically sincere ones, or that reports made ably and confidently are typically confident ones.’ [Lackey 2006a: 173]

However, it is not only reductionists and hybrid theorists who recognize a requirement that recipients of testimony assess the sincerity of testifiers. Indeed, a group of antireductionists, including Angus Ross [1986], Edward Hinchman [2005],5 and Richard Moran [2006]—a group whose trustworthiness.’ [M. Fricker 1998: 163] Though, in her recent [2007], M. Fricker is sensitive to the empirical evidence on the many ways in which prejudice can lead hearers astray in their assessments of their interlocutors’ trustworthiness or competence, M. Fricker nevertheless retains reliable monitoring of trustworthiness and competence as a condition on epistemically ‘virtuous’ testimonial reception.

4 Note also particularly Faulkner [2000: 599], where Faulkner writes that, ‘according to the hybrid theory, the acquisition of knowledge requires the audience possess some justification for acceptance.’ Later in the same paragraph, Faulkner suggests that, ‘if the audience’s justification for acceptance is no more than a judgment of credibility, then, ordinarily, it should provide no supplementary reason for the truth of the proposition accepted,’ implying that a judgment of credibility is the bare minimum that the hybrid view requires. [Cf. Faulkner 2000: 599-600]

5 Edward Hinchman [2005: 587]; in cases that don’t involve telling, but ‘merely hearing a speaker assert that p … [he] agree[s] with the reductionist
members David Owens, in a recent discussion,\(^6\) has termed assurance theorists—also seem to rely on recognitional factors in verbal communication that they take to be necessary conditions for testimonial justification. Thus, Edward Hinchman, like other assurance theorists, makes much of a putative parallel between assertion and promising. Hinchman ties his account of testimonial justification to a form of speech-act he terms telling, one component of which is that a recipient of a telling recognize the teller’s intention ‘that [the recipient] gain access to a prima facie entitlement to believe that p.’ [Hinchman 2005: 587]\(^7\)

Even anti-reductionists who would not count themselves among the assurance theorists have been tempted to embrace the notion that licensed acceptance of testimonial evidence requires monitoring for the trustworthiness of the testifier on the part of the recipient of that testimony. Thus, Goldberg [2007] has recently argued that his preferred anti-reductionist account of testimonial warrant allows us to grant that ‘some “monitoring for trustworthiness” is a necessary condition on the justified acceptance of testimony,’ noting that ‘there would appear to be something wrong with an account of justified acceptance on which a gullible hearer counts as justified … under conditions in which, had there been defeaters whose presence would have prompted any reasonable adult … not to accept the testimony, she would have accepted it.’ [Goldberg 2007: 162]\(^8\)

It is important at this point to clarify some of the terms that we will employ in the subsequent discussion. In particular, it will be useful to provide

\(^6\) David Owens [2006: 115-9]; Owens explicitly cites Ross and Moran, though not Hinchman.

\(^7\) Similarly, Moran suggests that ‘the speaker’s intent … is that for the audience the very fact that this speaker is freely and explicitly presenting P as worthy of belief constitutes his speech as reason to believe that P,’ and that, ‘of course, … the right conditions for such an act (e.g., that he possesses the relevant knowledge, trustworthiness, and reliability) … can themselves be construed as evidential.’ [Moran 2006: 289]

\(^8\) Of course, Goldberg’s nod to anti-individualism is that the ‘monitoring’ that is required would be allowed to be fully sub-doxastic, unavailable to the awareness of the agent, whereas—as we saw above—Fricker’s requirement was that the agent be able to recognize a testifier’s trustworthiness, a recognition that, presumably, requires that the agent be at least potentially aware of the evidence for that testifier’s trustworthiness. Where all of the parties so far discussed in this section agree, however, is that there ought to be a necessary condition on licensed acceptance of testimonial evidence that the recipient of testimony ‘monitor for trustworthiness’.
definitions of the notions of competence, trustworthiness, and monitoring. For our purposes, an informant will be said to be competent with respect to the question as to \( p \) just in case the informant’s testimony that \( p \) is reliable or otherwise truth-conducive.\(^9\) An informant will be said to be trustworthy with respect to the question as to \( p \) just in case the informant wouldn’t testify that \( p \) if she didn’t believe that \( p \).\(^{10}\)

One way to attempt to distinguish between—at least some—reductionists and anti-reductionists would be by considering their epistemological externalist credentials—viz., whether they require that a subject’s assessment of a testifier’s credibility be available to the awareness of the subject—rather than whether or not they hold that monitoring the trustworthiness of one’s interlocutor is a necessary condition for testimonial warrant. Since, however, we are here interested in the commonalities among all of those theorists—reductionists, anti-reductionists, and hybrid theorists—who hold that such monitoring is a necessary condition, in what follows we will use the term ‘monitor’ in a way intended to be noncommittal with respect to whether the monitoring in question involves evidence that is (at least potentially) available to the awareness of the agent.

Even so, there are a number of ways left open to us to understand the notion of monitoring under consideration here. For the purposes of the following discussion, we will focus on two particularly significant forms of monitoring trustworthiness of one’s interlocutors. In analogy to the twin epistemic goals of maximizing true beliefs and minimizing false beliefs, we will consider the two monitoring strategies of (1) adducing positive evidence that one’s interlocutor is trustworthy or competent and (2) exercising a sensitivity to evidence that one’s interlocutor is deceptive as two distinct processes. Indeed, there is at least some empirical evidence to suggest that subjects’ deployment of these two strategies is not the function of one mechanism, but rather that subjects employ separate strategies for adducing positive evidence of trustworthiness (or competence) and for exercising a sensitivity to deception. [Ekman et al. 1999: 265]

\(^9\) In borrowing the formulation ‘reliable or otherwise truth-conducive’ from Lackey [2006b: 437], I am intentionally remaining neutral on the question of the specific way in which the informant’s testimony is to be certified as truth-conducive—i.e., whether the reader prefers a sensitivity, safety, proper-functional, or other sort of account.

\(^{10}\) Note that, on the meanings of ‘competent’ and ‘trustworthy’ sketched here, it may be sufficient for one to acquire testimonial warrant that one’s informant merely be competent—and not trustworthy. (See, e.g., the discussion in Lackey [2006b] for thought experiments intended to motivate this view.) Since the necessary condition considered here—e.g., as laid out in (1.2) in section 1 above—is a disjunctive one, we need not concern ourselves further with this question here.
3. Evidence for the Unreliability of Monitoring: The Unreliability of Judgments of Trustworthiness

As we noted above, one source of evidence that could play a role in a subject’s licensed acceptance of testimony would concern positive judgments of their interlocutors’ trustworthiness. Positive judgments of trustworthiness depend for their reliability on a subject’s sensitivity to their interlocutors’ traits that reliably signal their trustworthiness. Unfortunately, however, a growing amount of social psychological literature on trustworthiness and deceit would suggest at least two main sources of difficulty for any account that has as a component a necessary condition involving sensitivity either to trustworthiness or to deceitfulness. The first is that there may well be no uniform, stable set—or sets—of traits signalling trustworthiness or deceitfulness. The second main source of difficulty is that, even if there are traits signalling trustworthiness or deceitfulness, subjects aren’t reliably sensitive to those traits. We will briefly consider the first source of difficulty, before turning to the question of subjects’ sensitivity to trustworthiness, in the remainder of this section, and to deceitfulness, in the following section.

Though it has not yet been the primary focus of investigation, many who have studied issues of deceitfulness and trustworthiness have questioned whether there are in fact any uniform, stable set(s) of traits signalling trustworthiness or deceitfulness. Thus, Zuckerman et al. [1981] refer to the widely accepted understanding that there is no one behaviour or set of behaviours that reliably co-occurs with lying. Similarly, Buller and Burgoon [1996] note that, given that patterns of behaviour depend on a deceiver’s expectations, goals, motivations, and degree of familiarity with their interlocutor, as well as with an interlocutor’s degree of suspiciousness, there is unlikely to be one profile of deceptive behaviour. Without such a profile, or profiles, of behaviours, however, there would of course be no question of subjects’ being able to be sensitive to the trustworthiness or deceitfulness of their interlocutors, as there would be no indicators to which the subjects could be sensitive in the first place.

For the discussion in the remainder of this section and in the section that follows, however, we will proceed as if there were such stable sets of traits indicating trustworthiness and deceit. We will suggest, though, that even if there are such traits, there is no reason to believe that subjects are reliably sensitive to those traits. Indeed, the preponderance of evidence from the social psychological literature is that we are terribly unreliable at responding to indicators of trustworthiness or deceit. Thus, though Stiff et al. [1989] note that ‘researchers have been less concerned with identifying the relation between cues related to actual truth and deception and those related to judgments of honesty and deceit,’ they do note that, of the four studies they found that reported on both sets of cues, ‘in all four studies, cues related to actual truth and deception differed from those related to judgments of honesty and deceit.’ [Stiff et al. 1989: 555; my italics]
This result, of course, will come as little surprise to anyone who has even cursorily reviewed the research in social psychology on influence and belief. For example, one of the classic studies, Dion et al. [1972], demonstrated that subjects stereotype based on physical attractiveness, attributing more socially desirable traits to those who are more attractive. And, as Chaiken [1979] demonstrated, the influence of attractiveness extends to trustworthiness as well; in her findings, attractive communicators were more likely to be believed than unattractive ones. Furthermore, Chaiken and Eagly [1983] showed that there is a similar link—perhaps not unrelated—between the likeability of testifiers and their effectiveness at persuading their interlocutors. There is little reason, however, to think that subjects relying on attractiveness or likeability are in fact responding sensitively to actual trustworthiness.

To take another case, consider the suggestion by Lackey, quoted in section 2 above, that ‘reports made with sustained eye contact are typically sincere ones, or that reports made ably and confidently are typically confident ones.’ To the extent that Lackey’s suggestion was one about our intuitive judgments concerning signs of trustworthiness, it is accurate: we do in fact take sustained eye contact to be an indicator of sincerity. [Cf. Claiborn 1979, Hensley and Doob 1978] However, though we rely on eye contact as an indicator of sincerity, taking sustained eye contact to be a sign of trustworthiness, this is in fact problematic, as eye contact has consistently been shown not to be a reliable indicator of sincerity. Indeed, as Sitton and Griffin [1981] demonstrate, deceivers in fact maintain more sustained levels of eye contact than truth tellers. Another intuitively held belief regarding sincere interlocutors is that they have no problem ‘keeping a straight face’ when communicating, but evidence suggests that liars produce no more nervous smiles than sincere interlocutors; in fact, studies suggest that liars smile somewhat less often. [Bond et al. 1985]\(^{11}\)

4. Evidence for the Unreliability of Monitoring: The Unreliability of Deception Detection

Given the empirical evidence regarding our unreliability at detecting trustworthiness, one might immediately conclude that we are equally bad at detecting deception. This, however, would be too quick. For, as we noted in section 2, there is some empirical evidence to suggest that the mechanisms employed by subjects to judge trustworthiness and to detect deception are in fact distinct. Indeed, Ekman et al. [1999] suggest that ‘the facial, bodily, and vocal differences between lying and truthfulness’ indicate that ‘there may be

\(^{11}\) Nor are education and expertise effective inoculations against the effects of charisma, appearance and presentation. One often-cited demonstration of this is the discussion of Dr. Myron Fox in Naftulin et al. [1973]. For more on monitoring of competence, see the discussion in section 6 below.
more identifiable signs of lying than of truthfulness.' [Ekman et al. 1999: 265] This would suggest that we must separately consider the reliability of strategies of deception detection from those of adducing positive evidence of trustworthiness.

Unfortunately, there is no empirical support for the notion that subjects are reliable detection detectors. Rather, as Ekman and O’Sullivan [1991] note, ‘in every study reported, people have not been very accurate in judging when someone is lying. … Average accuracy in detecting deceit has rarely been above 60% (with chance being 50%), and some groups have done worse than chance.’ [Ekman and O’Sullivan 1991: 913] These results have been demonstrated even for professionals whose jobs presumably depend on their ability to detect deception—customs officials [Kraut and Poe 1980], federal law enforcement officers [DePaulo and Fife 1986], and police officers [Kohnken 1987]. Apparently, the ability to deceive without being detected is acquired early, as recent studies have shown that adults are no better at detecting deception in children than they are at detecting deception in adult subjects. [Crossman and Lewis 2006]

Even the few research results that have found some groups capable of detecting deception at a rate better than chance serve in fact to underscore the challenges involved in reliably responding to lying. Thus, though recently Ekman et al. [1999] were able to demonstrate that two law-enforcement groups and a group of clinical psychologists were highly accurate in detecting deception, the researchers had pre-selected the video samples so that the test cases would display, in the researchers’ words, ‘significant behavioral differences between the subjects who lied and those who told the truth.’ [Ekman et al. 1999: 263] This, of course, is a significant modification of the test conditions, considering—as we noted in section 4 above—that many instances of deception do not evidence any standard set of behavioural patterns. Even given this drastic method of pre-screening the test cases to create the most favourable conditions for the test subjects, however, the results were equivocal enough that the researchers concluded that ‘it is unlikely that judging deception from demeanor will ever be sufficiently accurate to be admissible in the courtroom,’ and that ‘most of us would do well to entertain some skepticism about our ability to detect deception from demeanor.’ [Ekman et al. 1999: 265]

Of course, one might argue that the studies referenced above dealt with deception detection in interlocutors with whom subjects had no prior dealings. Perhaps, one might suggest, subjects’ abilities at detecting deception improve over repeated exposure to—and increased familiarity with—particular interlocutors. There is, for example, evidence that experience increases accuracy of lie detection—as in Bond et al. [1990], where it was shown that Americans detected lies told by other Americans 55% of the time, but were no better than chance at detecting lies told by Jordanians, whereas Jordanians detected lies told by other Jordanians 57% of the time, but detected lies told by Americans at no better than chance.
Furthermore, Zuckerman et al. \[1984\] demonstrate that greater experience with videotaped messages from senders delivering truthful and deceptive messages, coupled with feedback to subjects concerning whether they correctly identified the instances in which those senders had lied, was able to increase subjects’ accuracy at deception detection to between 60% and 70%.

Let us leave aside the—very real—concerns that this focus on communicative interactions with interlocutors with whom subjects are familiar doesn’t speak at all to the wealth of information subjects receive from their interactions with relative strangers. Even limiting our consideration to the possibility of deception detection in interactions with subjects with whom we are familiar, there is call for extreme caution in pursuing this line of response. The difficulty with the Bond et al. \[1990\] and Zuckerman et al. \[1984\] experiments is that they involve the examination of videotape evidence, thereby not correcting for the fact that, in real interpersonal interactions, just as potential detectors of deception gain experience from repeated exposure to their interlocutors, potential deceivers should also become more adept at tailoring their performance to their audience. Thus, Buller and Burgoon \[1996\] suggest that ‘deceivers in interactive contexts should display increasing immediacy and involvement, pleasantness, composure, fluency, and smooth turn taking over the course of the interaction.’ \[Buller and Burgoon 1996: 220\]

This would seem to be borne out by the evidence. For example, as McCornack and Parks \[1986\] demonstrate, though intimate partners have greater confidence in their ability to detect deception, they are in fact poorer at detecting deception in their partners than in other interlocutors. Indeed, research by Dunbar et al. \[2003\] suggests that interactive deception differs from noninteractive deception and that the interactive situation gives the upper hand to deceptive communicators rather than to deception detectors. \[Cf. Millar and Millar 1995\]

5. Where Do We Go From Here? Epistemological Implications of the Empirical Evidence

Recall the inconsistent triad with which we began our discussion:

1.1 The overwhelming majority of subjects are not reliable in monitoring their interlocutors for trustworthiness, deceit, and competence.

1.2 A necessary condition for acquiring warrant from testimony is that subjects reliably monitor their interlocutors for trustworthiness, deceit, or competence.

1.3 The overwhelming majority of subjects do, at least occasionally, acquire warrant from testimony.
As our survey of the empirical literature has demonstrated, neither of the monitoring strategies, i.e., neither judgments of trustworthiness nor deception detection, yield reliable information in most subjects. Thus, we have strong evidence in favour of (1.1)—at least with respect to subjects’ unreliability in the case of monitoring for trustworthiness and deceit. Thus, for those attempting to cling to (1.2), two options remain: (a) we can abandon (1.3), thereby embracing scepticism, or we can further explore ways in which (b) one might yet find a way of reconciling (1.2) and (1.3) with the empirical data—perhaps, in particular, by focusing on the as-yet undiscovered possibility of monitoring for competence. We will briefly discuss the sceptical option before turning to a discussion of the reconciliation option in the remainder of the paper.

If one is convinced by the suggestion—accepted, as we saw in section 2, by a wide swathe of those theorists writing on testimonial warrant—that licensed acceptance of testimony requires monitoring on the part of the recipient of that testimony, whether for trustworthiness or deception or both, then, given the empirical evidence adduced in sections 3 and 4 concerning the lack of reliability of such monitoring in human subjects, it would seem that the only conclusion is a healthy scepticism regarding the status of testimony as a source of epistemic warrant. However, it is almost universally accepted that testimony is a source of epistemic warrant. In the current state of the debate, scepticism regarding testimonial evidence simply isn’t a live option.\footnote{For a nuanced, dissenting discussion that begins from the unquestioned assumption that testimony is an indispensable source of belief, but that then proceeds to question whether it is a source of evidence, cf. Fumerton \cite{2006}.}

In what remains, let us turn to a consideration of ways in which reductionists or hybrid theorists might respond to the empirical data in a way that would allow them to reconcile theses (1.2) and (1.3) with the data adduced in support of (1.1) so far.\footnote{Although, for interests of space, the discussion here focuses on reductionists and hybrid theorists, there are reasons to think that the empirical results ought to be troubling to many externalist anti-reductionists as well. (Indeed, e.g., Goldberg, in conversation, has admitted as much.) The particular difficulty for the externalist is that it would seem that—\textit{ex hypothesi}, given that she is here supposing that the unreliable traits of the testifier are causally sufficient to induce belief—the recipient’s belief in the testimony is, in fact, \textit{based} on the unreliable traits of the testifier, viz., attractiveness, likeability, etc. That is, what explains why it is that the recipient believes the testimony is that the testifier is attractive, or likeable, etc., rather than that the testifier is reliable.

Indeed, even were it possible for the externalist to redescribe the case as an instance of a reliable process, the most natural understanding of that redescribed case would be to see it as an instance of a reliable method—believing a reliable informant—adopted on the basis of the application of an}{-11-}
forms of modified reductionism that survive the challenge of the empirical evidence telling against the reliability of monitoring for trustworthiness or deception: (i) modified monitoring, which abandons the requirement that subjects monitor for trustworthiness/deception in favour of a requirement merely that subjects monitor for competence or (ii) intellectualism, which abandons all monitoring requirements in favour of a requirement that recipients base their acceptance of testimonial evidence on deductive or inductive arguments in support of that evidence.

The difficulty with modified monitoring is that it is seldom possible for subjects to monitor for competence directly. Rather, subjects generally monitor for competence by monitoring for the trappings of competence—professional titles, clothing, or the fruits of success. Such monitoring by itself, however, as a function of what Cialdini [2007] has dubbed the ‘click, whirl mode’ of reacting, is about the symbols of authority, rather than the substance. Examples of the unreliability of such monitoring, though, abound in the social psychological literature.

Thus, to cite a classic case, Hofling et al. [1966] conducted a study in which, on separate occasions, a total of twenty-two different nurses’ stations on various wards of a hospital received identical phone calls in which a researcher identified himself as a hospital physician, unknown to the nurse on call, and directed the answering nurse to give an obviously and dangerously excessive dosage of an unauthorized medication to a specific ward patient. Researchers were testing the nurses’ reactions to four elements of the situation that should have roused their suspicions: (1) transmitting a

unreliable one—believing whoever is attractive/likeable. On many externalist accounts, however, beliefs deriving from such a method, arrived at in such a manner, would not ultimately be justified. Compare Goldman [1986], commenting on an analogous case: ‘Pending specification of some legitimate (metareliable) psychological process, one is left with the suspicion that … the believer has simply [by chance] stumbled upon a method which … is [reliable]. In that case the belief is not justified, on my theory.’ [Goldman 1986: 110]

Nor are other popular externalist accounts, such as proper functionalism [cf. Plantinga 1993] or virtue epistemology [cf. Sosa 1991, Sosa 2007], likely to have an easier time with this example. For, despite the fact that, if the experimental data is correct, believing whoever is attractive/likeable is in fact an instance of S’s ‘cognitive faculties … functioning properly’ and S’s ‘cognitive environment is sufficiently similar to the one for which [her] faculties are designed,’ it is unlikely that ‘the design plan governing the production of [S’s] belief … involves, as purpose or function, the production of true beliefs,’ as would be required by Plantinga’s proper functionalism. [Plantinga 1993: 194] Nor, as Sosa’s account would require, would it be easy to see a method like believing whoever is attractive/likeable as a 'manifestation of epistemic virtue.' [Cf. Sosa 2007: 23]
prescription by phone was in violation of hospital policy, (2) the medicine prescribed had not been cleared for use by the hospital in question, (3) the dosage prescribed dangerously exceeded recommended usage, and (4) the prescription was ordered by a person the nurse had never met, seen, or spoken to on the phone. Nevertheless, despite all of these indications that something in the situation was amiss, 95% of the nurses studied readied the doses and were prepared to administer the drugs—before being stopped by a secret observer, who only then revealed to them the nature of the experiment. Nor are subjects generally aware of the extent to which such automatic responses to the trappings of authority have power over them: when, in the same study, a separate group of thirty-three nurses and student nurses were asked how they would have responded in the experimental situation, only two admitted that they would have administered the medication as ordered.

To cite a further example, Cooper et al. [1996] report that a mock trial jury was far more likely to be convinced by jargon-filled, incomprehensible testimony from an expert witness introduced as having outstanding credentials than they were by the same testimony from the same witness introduced as having shaky credentials. Indeed, the same study showed that, when the witness was introduced as having outstanding credentials, the jury found him almost twice as persuasive when presenting his argument incomprehensibly than when he presented the same testimony using terms that the layperson could follow—that is, the jury found the expert more convincing, far more convincing, when they didn’t understand what he was talking about. Cooper and his colleagues concluded that the jury preferred using the shortcut of relying on the putative expert’s credentials to having to evaluate his arguments on its merits—that is, they preferred the ‘click, whirr’ mode to actually having to assess the testifier and his testimony.

The literature is replete with similar cases. Bickman [1974] demonstrated that, in an identical communicative interaction, subjects find interlocutors more persuasive when those interlocutors are dressed in a security guard’s uniform than when they are dressed in normal attire—by a difference in compliance rate of 92% to 42%. Nor need a uniform compel acceptance; as Lefkowitz, et al. [1955] demonstrated, a man in a business suit and tie was far more persuasive—by a more than three-to-one margin in compliance rate—than the same man in a work shirt and pants.

Of course—and in contrast to the empirical evidence marshalled in sections 3 and 4—the evidence reviewed here does not conclusively demonstrate that subjects aren’t normally reliable in detecting expertise. For, even if those subjects are in fact merely sensitive to the trappings of expertise, rather than expertise itself, if it is the case that those trappings are reliably connected to the presence of expertise, then subjects sensitive to the

---

14 Bickman [1974] involved male interlocutors; Bushman [1988] achieved similar results using females posing as either security guards or bystanders.
trappings will, thereby, be reliably sensitive to expertise. Even so, the evidence presented here does suggest that the modified monitoring reductionist position is unstable, in the following sense.

The evidence presented here suggests that, even if one were to grant that subjects are sensitive to expertise, then it is only because they are sensitive to the outward signs of expertise, and because those outward signs are in fact a reliable indicator of expertise. Now, however, the question arises: need subjects be aware of the fact that those outward signs are reliable indicators of expertise? An affirmative answer to this question suggests that this form of modified monitoring reductionism is in fact a form of intellectualist reductionism, to which we will turn in section 6. A negative answer to this question, however, suggests that this form of modified monitoring is not in fact reductionist at all; indeed, it would be difficult to maintain this position against full-fledged externalist anti-reductionism.

6. The Intellectualist Response

Intellectualist reductionism eschews monitoring requirements in favour of requirements that recipients of testimony rehearse arguments providing deductive or inductive support for their acceptance of testimony as evidence. Of course, intellectualist reductionism might plausibly require that subjects be sensitive to features of the communicative situation likely to be salient with respect to such arguments; what distinguishes intellectualism from monitoring is that intellectualism requires that subjects engage in an explicit inference, whereas monitoring—even of the reductionist variety—would seem to connote more of the ‘click, whirr’ mechanism of which Cialdini writes.

Thus, to take a particular example, an intellectualist might require that subjects reason as follows: This woman in my hospital room is wearing a white lab coat, has a stethoscope draped around her neck, and is holding an official-looking clipboard; people with those outfits are likely to be doctors; doctors are likely to know what they are talking about with respect to medical conditions; thus, I ought to take her statement that my medical condition is p and that the appropriate courses of treatments for that condition are q₁, …, qₙ as evidence that my medical condition is p and that the appropriate courses of treatments for that condition are q₁, …, qₙ. Such a view has the advantage that, if the premises are true, subjects employing reasoning as in the example here will be able to be seen to be reasoning properly even in those—presumably sufficiently rare—fraudulent cases in which someone might be exploiting the trappings of expertise for her own nefarious ends. For those for whom a subject’s living up to her epistemic

However, evidence suggesting that subjects are unreliable in source monitoring tasks, discussed in sections 6 and 7 below, would seem to caution against excessive optimism even with respect to the question as to whether subjects are in fact reliable in detecting expertise.
responsibility is of primary concern, intellectualist reductionism may well be an attractive response to the empirical evidence adduced in sections 3 and 4 and in section 5 above.

Of course, those following the debate between reductionists and anti-reductionists in the epistemology of testimony will recall that the whole point of the recent move to monitoring positions in reductionism was to escape the implausible requirement, entailed by intellectualist reductionism, that subjects explicitly rehearse arguments in favour of their acceptance of instances of testimonial evidence. This requirement seems implausible in that it seems at odds with the phenomenology of belief formation on the basis of testimonial evidence. The monitoring vs. intellectualist dichotomy closely parallels a dichotomy between what Cacioppo and Petty [1979] have termed ‘central’ and ‘peripheral’ routes to attitude formation. Employing the ‘central route’ involves a concerted effort to investigate the facts of a case, whereas the ‘peripheral route’ involves relying on shortcuts—emotional reactions to interlocutors, signs of their putative expertise—to ease our decision-making.

As much of the empirical evidence we’ve reviewed here suggests, however, it would seem that subjects often employ peripheral routes in acquiring beliefs on the basis of testimony. Indeed, recall the lesson of Cooper et al. [1996]: subjects prefer to employ peripheral routes and are often more convinced when forming attitudes via peripheral routes than when forming them through central routes. Given this, it would seem that the intellectualist could only maintain a requirement of inferential support for testimony at the cost of a very significant scepticism about the extent of warranted, testimonially derived beliefs. Since most testimonially derived beliefs, as the experimental data show, are formed automatically, via peripheral or ‘click, whirr’ routes of attitude formation, they would lie outside the pale of inferentially licensed testimonial beliefs, and would thus be, for the intellectualist, unjustified.

At this point, however, it would be open to the intellectualist to appeal to the distinction between the acquisition of a belief and its justification. For, even if a subject acquired a belief in an improper manner, as long as there are strong reasons sustaining the subject’s confidence in that belief, then the belief may yet count as justified for that subject. Thus, even if a subject initially acquired a belief via one of the unreliable mechanisms considered above—testimony of a physically attractive source, say—the subject could nevertheless be justified in the belief if she subsequently gained further, reliable, evidence in support of the belief, evidence that now is sufficient to sustain her belief. Indeed, it is this sort of support to which a subject would appeal in justifying her belief to others—not, ‘This good-looking person told me,’ but ‘I read it in the New York Times.’

Unfortunately, the empirical evidence suggests that this route is not in fact a promising one for the intellectualist. For such justifications depend upon a subject’s ability to identify introspectively the sources of their beliefs and, in the case of retrospective analysis of beliefs previously acquired, to
recall the sources for information retrieved from memory, in order then to assess its quality or to muster it in service of justification. There is a great deal of empirical evidence, however, to suggest that subjects are highly unreliable both at identifying the sources of their beliefs introspectively and at correctly recalling the sources of information retrieved from memory.

As Kornblith [2002] has persuasively argued, subjects are abject failures at introspectively identifying the actual sources of their beliefs. [Cf. Kornblith 2002, Chap. Four] Nor, indeed, are subjects any better at the retrospective analysis of beliefs previously acquired. There is a vast and recent psychological literature regarding the processes by which subjects form judgments regarding the processes by which subjects form judgments regarding support for information retrieved from memory, processes that psychologists discuss under the heading of source monitoring. [Cf. Johnson et al. 1993] As psychologists have demonstrated, subjects are prone to widespread and systematic source monitoring failures.

Of particular relevance to the discussion here are the studies in Fragale and Heath [2004]. Fragale and Heath demonstrated that subjects often rely on a ‘belief force equals credible source’ heuristic—assuming that statements that they believe originate from credible sources—to make judgments concerning the sources of their beliefs. For example, in one of the studies, Fragale and Heath manipulated subjects’ beliefs in a variety of food contamination allegations by varying the number of times subjects were exposed to those allegations, taking advantage of the well-established effect through which repeated exposure to statements is sufficient to induce belief in those statements. [Cf. Arkes et al. 1989, Hawkins et al. 2001] After this, they then questioned subjects as to what the likely source of those allegations was, either the National Enquirer or Consumer Reports. As predicted, subjects with greater exposure to the allegations were far more likely to identify Consumer Reports as their source. In summary, Fragale and Heath conclude that their ‘studies suggest that informational credentials may be generated more spontaneously … than previously thought.’ [Fragale and Heath 2004: 234]

Thus, the evidence would suggest not only that subjects are poor perceivers of the sources of their beliefs but also that they are inaccurate at retrospectively analyzing those sources. Given this, however, an intellectualist position that abandoned an emphasis on central—i.e., inferential—routes of belief formation via testimony in favour of an emphasis on placing the responsibility on subjects to assess the evidence sustaining their belief would seem to have no greater prospects for success.


At this point in the debate, some defenders of the monitoring condition might be tempted to suggest that perhaps our reading of (1.2) misconstrues the force of the necessity that recipients of testimony monitor
their interlocutors. Such defenders of a monitoring requirement might suggest that reliable monitoring need not prove to be an actually prevalent epistemic practice, but rather that reliable monitoring is instead an ideal of good epistemic practice, a regulative epistemic norm. In order to embrace monitoring as such a regulative norm, this line of rejoinder continues, it is enough to show that, with sufficient experience, a reasonably capable epistemic agent could act in accordance with the monitoring requirement.\footnote{Indeed, Lizzie Fricker has suggested this argumentative move to me, in conversation, as has an anonymous referee for this journal. A still different, though related, strategy, suggested to me by an anonymous referee for this journal, would be to grant that monitoring has no reliable truth connection, but to maintain that it is nevertheless an epistemic duty. Without entering into a consideration of whether there are any properly epistemic, though non-truth-related, goals, I will limit myself here to a discussion of truth-related epistemic goals.}

The difficulty with this move is that, again, it runs up against very strong empirical evidence against the idea that experience improves one’s ability to detect trustworthiness, deception, or competence. Thus, e.g., Garb’s [1989] review of studies measuring the effect of experience on clinical psychologists’ ability accurately to make clinical judgments catalogues the wealth of data demonstrating that experience yields no such accuracy increase. Indeed, in the case of clinicians’ judgments the track-record is so dire that Dawes was confident enough flatly to state that ‘the empirical bottom line has already been established: Appeals to experience per se are invalid because experience per se does nothing to enhance accuracy.’ [Dawes 1994: 109]

If experience cannot even increase clinical psychologists’ accuracy in diagnosing psychological conditions, it would seem bizarre to suppose that experience would increase laypersons’ accuracy in detecting trustworthiness, deception, or competence. In the case of trustworthiness- and deception-detection, Kohnken’s [1987] study demonstrated that even experienced police officers who received specific training to improve their monitoring capabilities were no better than chance in trustworthiness- or deception-detection. Furthermore, a study by Brandt et al. [1980] demonstrated that, though increasing the number of iterated trials with the same interlocutor pair initially increased the accuracy of judgments of trustworthiness or deception, with continued trials recipients’ accuracy then declined. Brandt and his colleagues suggest that, over time, recipients gain too much information about their interlocutors, and become unable to filter out the probative information from the noise.

Indeed, the fact that experience should fail to increase monitoring accuracy should be unremarkable, when one recalls the brief discussion at the conclusion of section 6 of the prevalence of systematic source monitoring failures among subjects. If experience were to be helpful in improving
performance, presumably this would be the result of a subject’s ability accurately to recall the sources of their beliefs and accurately to track which of those beliefs were true. As the discussion of source monitoring suggests, however, subjects are accurate in neither of these two tasks. Thus, e.g., Dawes [1994] traces the failure of experience to improve performance in clinicians’ judgments essentially to failures of source monitoring. [Dawes 1994: 109 - 121]

Of particular significance for our discussion here, however, is that, though experience does not increase accuracy, it does increase subjects’ confidence in the accuracy of their judgments. Indeed, Kohnken’s [1987] study cited above noted an inverse correlation between accuracy and confidence: the less accurate Kohnken’s subjects were, the more confident they were in their assessments of trustworthiness or deception. This would perhaps explain the move to an appeal to a monitoring condition as a regulative epistemic norm, since we all like to think that, at least with experience, we could improve our monitoring abilities. The evidence, however, would not seem to speak well of our chances.

Though the role of experience in improving competence monitoring is more difficult to study, there is at least some evidence to suggest that competence monitoring is no more amenable to improvement on the basis of experience. Thus, a number of studies in college and medical school admissions have demonstrated that there is no positive correlation between interviewer experience and validity in judgments of applicant ability. [Cf. Dawes 1994: 86 – 89] Furthermore, as noted in section 6 above, the failure of subjects at source monitoring tasks is equally relevant for the question of competence monitoring. Given this fact, there is little reason for optimism here as well.

In closing, note that, at the very least, even this strategy on the part of defenders of a monitoring requirement, rather than obviating the need to pay attention to the empirical data, in fact demands careful attention to the empirical data. The question here is not one of facts versus norms, or empirically based argumentation versus philosophical intuition, but one of evidence. We have, in sections 3 - 7, seen a wealth of evidence underwriting the claim, laid out in (1.1) that subjects are unreliable in their monitoring of trustworthiness, deception, and competence. The result is a tension between our unwillingness to embrace scepticism (1.3) and the widespread adherence, among epistemologists, to a monitoring requirement for testimonial warrant (1.2). Since that adherence to (1.2), however, has nothing more to support it than bare philosophical intuition, it would seem to be the weak link in the inconsistent triad.17

17 This is not to say, of course, that there will be no requirement to replace it, or that we should just license the acceptance of testimony by liars (to cite an example raised to me, in conversation, by Lizzie Fricker, and by an anonymous referee for this journal). Note, to briefly sketch merely one
8. Conclusion

Recent theories in the epistemology of testimony have been virtually univocal in their call for a requirement that recipients of testimony monitor testifiers for their credibility as a necessary condition on testimonial justification. Despite this near-unanimity, substantial empirical evidence, I have argued, favours the abandonment of such monitoring requirements. The links between recognizable properties of testifiers that reliably indicate them as sincere and competent and the beliefs formed by recipients of testimony emerge as much more tenuous and less capable of supporting such robust requirements as we might otherwise have thought. In this way, a certain common-sense conception of the responsibilities accruing to recipients of testimony actively to avoid gullibility would thereby seem to be at least potentially misleading. Thus, proper attention to the empirical evidence regarding the ways testifiers and their interlocutors actually interact would seem to undermine all those epistemologies of testimony that attempt to explain testimonial justification even in part in terms of the duties borne by recipients of testimony—including all forms of reductionism in the epistemology of testimony.

example, that a form of reliabilist externalism that did not have a monitoring requirement, but that conceived of processes as involving interlocutors—i.e., not just trusting, but trusting-\(T\), where \(T\) is a particular testifier—would, in cases where testifiers were liars, not license trusting liars. If I have understood him correctly, Goldberg’s [forthcoming] fleshes out one such version of a reliabilist externalism with regard to testimonial warrant. I develop my own such account in Shieber [2010b].
References


Owens, D. 2006. Testimony and Assertion, Philosophical Studies 130/1: 105-29.


Shieber, J. 2010b. Toward a Truly Social Epistemology: Babbage, the Division of Mental Labor, and the Possibility of Socially Distributed Warrant. Manuscript.


