

# Testimonial Contractarianism

## Abstract

According to anti-reductionism in the epistemology of testimony, testimonial entitlement is easy to come by: all you need to do is listen to what you are being told. Say you like anti-reductionism; one question that you will need to answer is how come testimonial entitlement comes so cheap; after all, people are free to lie.

This paper has two aims: first, it looks at the main anti-reductionist answers to this question and argues that they fail. Second, it goes on a rescue mission on behalf of anti-reductionism. I put forth a novel anti-reductionist account, which I dub ‘Testimonial Contractarianism’. According to the view defended here, in virtue of the social contract in play, compliance with the norms governing speech acts is the default position for speakers. Insofar as norm compliance is the default for speakers, I argue, all else equal, entitlement to believe is the default for hearers.

## 1. Introduction

We believe what other people tell us. I trust that you are on the way to the mall if you tell me you are on the way to the mall; I believe that the price of oil has gone down because I heard it on the news. In the summer I will go hiking in the Carpathians; I remember my geography teacher told us that wild life is gorgeous over there. And I believed him.

We believe what other people tell us, and that’s all right. We don’t need to check on the facts ourselves; nor is elaborated detective work needed to verify the credentials of the testifier. We can just go ahead and believe on mere say-so: it’s fine. Call this view *Anti-Reductionism* (AR) about testimonial entitlement.<sup>1</sup>

---

<sup>1</sup> I do not mean to suggest that anti-reductionisms and reductionisms in the epistemology of testimony are two uniform, clearly delineated bunches. To the contrary, champions of both views make very distinct claims, concerning very distinct issues related to testimonial entitlement. See Lackey (2008) for a very useful taxonomy. Following champions of the views I discuss here, however, for the purposes of this paper, I am going to focus on the particular difference between reductionism and anti-reductionism when it comes to how heavy an epistemic burden they lay on hearer’s shoulders in the testimonial exchange: while the boundaries between the two camps are, by no means, clear, reductionists tend to require hearers to have some variety of access to their reasons to trust their testifiers, while anti-reductionism tends to deny such access is necessary. All this is still pretty vague, but one useful way to see the distinction that we care about here is to think of Pritchard’s (2004) taxonomy, distinguishing between what he dubs Credulism, on one hand, and Reductionism on

Now, say Anti-Reductionism is true; one interesting philosophical question that readily arises is how come testimonial entitlement comes so cheap. After all, people are free to lie. Furthermore, they are rational to attend to their own interest, rather than someone else's. It makes sense for you to lie that you're going to the mall if it serves your interests to have me believe that that's the case. Why, then, can I trust that you're telling me the truth? Why am I entitled<sup>2</sup> to believe you? Let us dub this problem the *Source Problem* for testimonial entitlement (*SP*): it concerns the source of our warrant to believe what we are being told by other people, on mere say-so.

There are two main anti-reductionist responses to the Source Problem available in the literature. According to Strong Anti-Reductionism (SAR) – championed, most notably, by Taylor Burge – the source of testimonial entitlement is a priori. Since rational capacities have the function of generating truth, we are *prima facie* entitled to believe rational speakers on mere say-so. The hearer need not do any work.

More recently, other philosophers defend less ambitious solutions to SP: according to people like Sandy Goldberg (2010, 2014), Peter Graham (2010, 2012) and John Greco (Forthcoming, 2015), the source of testimonial entitlement is social: there are social grounds for us to trust what other people tell us. Since the proposed answer to SP is more modest, so is the corresponding anti-reductionist condition on testimonial entitlement (henceforth Modest Anti-Reductionism). On Goldberg's view, for instance, hearer's entitlement is restricted to beliefs that were generated by speaker's epistemically trustworthy processes to begin with. According to Graham and Greco, hearers need to do a bit of epistemic work - albeit not much - in order to be entitled to their testimonial beliefs: on Graham's view, they should actively watch for signs of untrustworthiness. Greco's account is even less ambitious than that: depending on the social roles of the actors involved in the testimonial exchange, there are, he argues, situations when positive reasons to trust one's speaker are required.

This paper assumes Anti-Reductionism is true, without further defence. The scope of the paper is limited to finding a novel answer to the Source Problem on its behalf.

---

the other. Champions of reductionism include Adler (1994), Audi (1997, 2004, 2006), Fricker (1995), Hume (1739), Lipton (1998), Lyons (1997). For defenses of anti-reductionist (credulist) views, see, e.g. (Author 2017b), (Burge 1993, 1997, 1999), (Coady 1973, 1992), (Goldberg 2006, 2010, 2014), (Goldman 1999), (Graham 2010, 2012, 2015), (Greco Forthcoming, 2015), (Green 2016), (Reid 1764). For hybrid views, see e.g. (Faulkner 2011), (Lackey 2003, 2008), (Pritchard 2004).

<sup>2</sup> For the purposes of this paper, I use justified, warranted and entitled interchangeably. Nothing hinges on this.

<sup>3</sup> Thanks for xxx for pointing this out.

<sup>4</sup> In personal conversation, Sandy Goldberg argued that what is important is the recipient's entitlement to presume same cognitive processing types. Accordingly, this entitlement is generated for ordinary human beings insofar as we co-evolved with our conspecifics, and evolved as well to be social creatures part of whose sociality lies in

More precisely, the paper has two aims. The first is negative: I will argue that, on closer scrutiny, we have no good reason to go weak (Section #2); in order to do this, I look at the motivations put forth by champions of Modest Anti-Reductionism in favour of weakening the anti-reductionist condition on testimonial entitlement, and argue that they fail; as it turns out, SAR does exceptionally well in accommodating the data put forth in support of MAR.

Further, then, I move on to discussing the most notable SAR solution to SP on the market, championed by Tyler Burge. I argue that the argument for an a priori source of testimonial entitlement fails to go through due to function-theoretic considerations (Section #3).

In contrast, my second aim is constructive. In #4 and #5, I will go on a rescue mission on behalf of Anti-Reductionism. First, I take a closer look at the Source Problem and, in line with Faulkner (2011), I argue that, on its most plausible reading, SP amounts to a problem of cooperation. Second, in response to this, I develop a novel strong anti-reductionist account of testimonial entitlement: Testimonial Contractarianism. I argue that, in virtue of the social contract we are part of, we are prima facie entitled to trust that other people will obey social norms; we don't need positive reasons to believe that they will. Since it is plausible that some epistemic norms – to wit, norms governing speech acts – are also social norms, it follows that in virtue of the social contract we are part of we are prima facie entitled to believe what other people tell us; we don't need positive reasons to trust speakers. The account put forth is strongly anti-reductionist, in that it argues that speakers are prima facie entitled to believe what they are being told. It is a social anti-reductionism, in that, contra Burge, it takes the source of such entitlement to be a posteriori, sourced in the social contract, rather than a priori.

## **2. No Reason to Go Weak**

The Source Problem is, essentially, a problem of easy entitlement; what the worry amounts to, at first glance, is the following: given that speakers can lie, believing anyone who asserts anything on mere say-so amounts to faulty gullibility: hearers should be epistemically pickier than that. In the light of the easy entitlement worry behind SP, in recent years, several philosophers have offered versions of AR that are proposing some – albeit little – limitations to hearer's entitlement.

### *2.1 Goldberg's Socially Extended Entitlement*

According to Sandy Goldberg, our entitlement to believe on mere say so is sourced in the epistemic quality of the belief acquisition processes that

our speaker's assertion is based on to begin with. Here is the thought: it looks as though we are epistemically entitled to rely on a particular category of belief generating processes, call them 'privileged' processes: perception, memory and inference are such privileged processes. Now, why should we think that we are only entitled to the outputs of *our* privileged processes and not to the outputs of other people's privileged processes too? After all, it does not look as though the epistemic quality of these processes, what makes them privileged, is in any way dependent of them being hosted by one particular agent or another. Here is Goldberg:

[W]hatever meta-epistemological story supports principles of epistemic entitlement will be indifferent to the distinction between the subject to whom the process belongs, and the subject who is relying on that process. Indeed, it is hard to fathom why the metaepistemology of entitlements should care about this distinction (2014, 166).

One answer to this question, of course, is suggested by the Source Problem: because people can lie. Our access to the outputs of their privileged processes is mediated, and can be obstructed by people's ability to lie.

Goldberg grants as much; according to him, however, this fails to satisfactorily answer his question. After all, the question was one concerning epistemic significance: why should we think that the live possibility of lying has such devastating normative import, i.e. it renders someone else's privileged processes completely void of epistemic strength for me? After all, my own processes can also fail me; the mere possibility of failure, however, does not rob them of their privileged epistemic status:

Just as one who relies on memory must guard against memory-produced confabulation, so too one who relies on another's assertion must guard against insincerity (as well as other sources of unreliability) in assertion. This consideration is the stuff of defeaters, not something that shows that the source itself cannot be appropriately relied upon (2014, 168).

As such, according to Goldberg, our entitlement to believe on mere say-so is sourced in the privileged status of the processes employed by the hearer: "If one observes and understands another speaker's assertion that *p*, there are no relevant defeaters, *and the speaker's assertion that p was itself the result of 'privileged' processes operating in her* (, henceforth 'Goldberg's condition'), then one is entitled to believe that *p*"(2014, 174).

Now, note that Goldberg's view places no special burden on hearer's shoulders in the testimonial exchange; in virtue of this, it is an anti-reductionism. It is a modest variety thereof, however, because not

just any reliance on a say-so will generate epistemic entitlement. It needs be that the assertion at stake is itself properly based, on the ‘privileged’ epistemic processes of the speaker. Beliefs based on lies or other types of epistemically dubious assertions cannot enjoy entitlement.

Here is the main worry for this view, however: it concerns the generative power of testimony; notably, work by Peter Graham (2006) and Jennifer Lackey (2008) supports the view that testimony can generate knowledge by generating entitlement. In particular, it looks as though weirdly formed assertions – i.e. assertions which are not based on ‘privileged’ processes operating in the speaker - can also be knowledge generators: take Bertha, Lackey’s consistent liar (2008). As a result of a brain lesion, Bertha is prone to lie concerning her perceptual experiences of wild animals. Whenever she sees a deer, she will report otherwise. Enter Dr. Jones, who identifies the lesion, and, since he cannot fix it, he creates a new one that cancels its effects: now, every time Bertha sees a deer, she believes it is a horse, and every time she believes there is a horse in front of her, she reports it is a deer.

Several epistemologists agree with Lackey that Bertha is perfectly capable of generating testimonial knowledge. Of course, if Lackey is right about Bertha, Goldberg’s view is in trouble: after all, the processes underlying Bertha’s assertions are clearly too weird to qualify as ‘privileged’ in the relevant sense. If knowledge is generated in the hearer, however, and since Goldberg accepts that knowing implies having entitlement to believe, it looks as though Goldberg’s necessity claim is called into question: testimony generates epistemic entitlement.

Goldberg is, of course, fully aware of this potential worry; what he does in response is to dismiss Bertha-type cases as cases of knowledge generation. Since the intuition is hardly uncontroversial in these cases, he argues, this will not be too big a bullet to bite.

Agreed. Arguably, however, Bertha is just a limit case of a more general worry for Goldberg’s condition; i.e., a distinctiveness worry: why should we think that only creatures that function biologically just like us – i.e., have the same belief generating processes as ourselves – can be entitlement generating testifiers? To see the worry, consider, first, a variation of another case by Jennifer Lackey:

ALIEN: Sam, an average human being, is taking a walk through the forest one sunny morning and, in the distance, he sees some weird creature drop a book. Judging by her strange physical appearance, Sam identifies the creature as an alien from another planet; he has no further knowledge about this particular type of alien or the planet from which she comes. Sam recovers the book that the alien dropped, opens it and reads: ‘The President will have blueberry muffins for breakfast tomorrow’. Sam thereby comes to believe that the President will have blueberry muffins for breakfast tomorrow.

Cases like ALIEN are put forth in the literature to support one variety or another of a reductionist condition on testimonial entitlement. The thought is that Sam is, intuitively, not entitled to believe that the president will have blueberry muffins for breakfast, in virtue of the fact that he has no information about the epistemic credentials of the testifier, no positive reasons to believe her.

Now, irrespectively of whether we buy the reported intuition, or the reductionist explanation of this data point, one thing seems clear: whether Sam is entitled to believe or not does not seem to hinge on the cognitive apparatus of the alien, i.e. on whether the alien is employing a process figuring on Goldberg's 'privileged' list or not. Goldberg is well aware of this problem. In response, to ALIEN-type cases, he appeals to evolutionary considerations:

[...] I want to dig in my heels: insofar as the cognitive-psychology of a species of creatures really is distinct from human psychology, to just that extent I deny that there is any default entitlement to rely on (tokens of) their (apparent) say-so. [...] As we did not evolve with the aliens, there is no shared species history, and so no default entitlement for us to rely on them (2014, 180).

There are two problems with Goldberg's reply, however: the first concerns intuitive adequacy. To see the issue, consider, first, a slightly modified version of the case, call it ALIEN\*, in which, completely unbeknownst to Sam, this particular variety of aliens has a cognitive apparatus that's very similar to ours. Surely, this will not make much difference for the intuition triggered by the case.

Conversely, consider:

ORDINARY SPEAKER: Sam, an average human being, is taking a walk through the forest one sunny morning and, in the distance, he sees the President's personal assistant drop a book. Sam recovers the book, opens it and reads: 'The President will have blueberry muffins for breakfast tomorrow'. Sam thereby comes to believe that the President will have blueberry muffins for breakfast tomorrow.

Now, my anti-reductionist intuition is that Sam is entitled to believe what he reads in the book. Reductionists, of course, are likely to disagree.<sup>3</sup> Not much hinges on the type of intuition, however. Say that, unbeknownst to everyone involved, the President's personal assistant is one of the few people affected by a harmless mutation that renders her belief forming processes completely different than ours. This does not

---

<sup>3</sup> Thanks for xxx for pointing this out.

seem to change the intuition in this case – be it for entitlement or against it.

If that is the case, however, that is, if my diagnosis of ALIEN\* and ORDINARY SPEAKER is correct, it looks as though whether the speaker employs a ‘privileged’ – in the relevant, Goldbergian sense – belief formation process or not is irrelevant to testimonial entitlement.

The second problem concerns the legitimacy of Goldberg’s appeal to a specific variety of evolutionary considerations, i.e. social evolutionary considerations, in response to this problem. Recall that Goldberg argues that the reason why we should think that we are not entitled to believe aliens, employing non-privileged belief-formation processes, but we are entitled to believe each other, is because ‘we have not evolved *with* (emphasis added) the aliens’ (2014, 180). While I am ready to grant Goldberg that evolving *with* the aliens would have made a difference, I deny that evolving *like* the aliens matters for testimonial entitlement. The former is a consideration pertaining to social evolution. The latter is a consideration pertaining to biological evolution. In imposing a condition pertaining to the type of belief generation process employed by the speaker, rather than pertaining to the social relations between the hearer and the speaker, Goldberg’s account can afford relying on the latter, but not the former. In a nutshell, I agree with Goldberg that whether the hearer and the speaker evolved *with* each other might matter for testimonial entitlement; as a matter of fact, the last section of this paper will make appeal to just such considerations to answer the Source Problem. What I do not agree with – and what the cases above are meant to disprove - is that the speaker and the hearer need have evolved *like* each other – i.e. use the same belief formation processes - for success in entitlement.<sup>4</sup>

---

<sup>4</sup> In personal conversation, Sandy Goldberg argued that what is important is the recipient's entitlement to presume same cognitive processing types. Accordingly, this entitlement is generated for ordinary human beings insofar as we co-evolved with our conspecifics, and evolved as well to be social creatures part of whose sociality lies in our dependence on them for information. In the case of an alien, the human isn't entitled to presume that (even if she is mistaken in supposing that she is; perhaps the alien is cleverly disguised as a human).

There is a lot to be said about this, and it will be impossible to do justice to this response here; at a first glance, however, I suspect Goldberg’s reply might drag Goldberg’s account into a strength dilemma. Here is how: if entitlement is a thick enough epistemic state to (at least) entail the possession of the relevant concepts, the account will face over-intellectualization worries. If in order to be entitled to believe that *p*, one first needs to be entitled to believe that the testifier employs the same cognitive processes as her, the account threatens to become too strong to accommodate unsophisticated cognizers. For instance, small children will not be entitled to believe that there’s milk in the fridge based on their mother’s testimony because they lack the concept of cognitive process, and thus are not entitled to believe that their mother employs the same cognitive processing types. On the other hand, if one makes the concept of entitlement as thin as to not even presuppose concept possession, it is not clear what work it can still do for the anti-reductionist; after all, the reductionist might well be willing to grant that this thin epistemic state is easily present.

## 2.2 Graham's Filtering Requirement

According to Peter Graham, we humans have internalized social norms (Graham (2010, 2015)). Furthermore, the motivation provided by internalized social norms frequently takes primacy over motivations that accord with our immediate interest. Speakers have internalized a social norm that prescribes telling the truth informatively, and the motivation provided by this social norm frequently overrides any motivation that rational choice theory predicts we should have: speakers will frequently tell the truth even when it is in their best interest not to (Graham (2010, 223), (2015, 256), (2012, 112)). As such, Graham argues, the threat posed by SP is, to a large extent, averted:

[...] internalized social norms explain behavior that otherwise would have seemed very puzzling, at least from the point of view of rational choice theory. We have an example of behavior that goes against what rational choice theory would have predicted in the first place. Participants “see” what they are “supposed” to do in such a situation in terms of internalized norms, and those internalized norms then explain, in large part, what they do (2012, 112).

Here is, then, a rough sketch of Peter Graham's MAR: the hearer need not do much: all there is needed for (*prima facie, pro tanto*) testimonial entitlement on the hearer's side is for her to form her beliefs via a properly functioning process of comprehension and filtering that has the function of reliably generating true beliefs. Comprehension will be in charge with uptake, while filtering has the job of detecting indications of untrustworthiness.

Since independent inductive reasons for trusting the speaker are not required, Graham's view qualifies as a version of AR. What makes the view modest is the fact that it imposes an active filtering demand on the hearers, which narrows the range of entitlement conferring testimonial exchanges. The thought here is that even if simply taking a speaker's word at face value is tantamount to an objectionable form of gullibility, doing so after having filtered for indications of untrustworthiness isn't.

Now, one question that arises about Graham's view is: why should this filtering requirement appease Reductionist worries? Why should we think it satisfactorily answers SP?

In a nutshell, Graham's answer is a reliabilist one. Although social norms internalization affects testimonial exchanges, such that people don't lie as often as it would suit them, testimony is still a considerably less reliable source of information than non-agent-mediated



sources, such as perception or inference.<sup>5</sup> The filtering requirement, then, is added in the case of testimony in order to reduce this difference in reliability, and turn testimony into an equally good entitlement-conferring source.

However, empirical results seem to fail to support Graham's motivation for going modest. A wide range of studies testing our capacities for deception recognition show that we are very bad at it: our prospects of getting it right barely surpass chance (e.g. (Kraut 1980), (Vrij 2000) and (Bond and DePaulo 2006)). To see just how well-established this result is in the relevant psychological literature, consider the following telling passage from Levine et al.: "the belief that deception detection accuracy rates are only slightly better than fifty-fifty is among the most well documented and commonly held conclusions in deception research." (1999, 126)<sup>6</sup>

Crucially, it is not hard to see that if these studies are right and we detect deception with an accuracy rate that is barely above chance, the differences in reliability between those who accept testimony without further filtering and those who do make the additional effort of filtering will be negligible.<sup>7</sup>

Of course, this is bad news for Graham. After all, his reason for adding a filtering requirement to AR was to reduce this difference by increasing the reliability of testimony. What the above considerations indicate is that filtering fails to deliver the goods. As a result, Graham's motivation for going modest fails.<sup>8</sup>

### 2.3 Greco's Social Roles

Just how far does the gullibility worry extend? It's hard to deny that simply taking the speaker's word at face value will amount to objectionable gullibility in some cases. Thus consider:

**Case 1.** An FBI agent questions a suspect in a murder mystery.

**Case 2.** A used car salesman tells you that the vehicle is in mint condition.

---

<sup>5</sup> Graham, personal communication.

<sup>6</sup> Could all these studies be flawed? Sure. But given that there is a general consensus in the relevant literature, it does not seem tendentious to work on the assumption that the consensus view is correct.

<sup>7</sup> See (Michaelian 2010) and (Faulkner 2011, 86-87) for a similar point. See Fricker (Forthcominga) for skepticism about the empirical results.

<sup>8</sup> One way to go in support of the Graham view at this stage is to argue that, even though we are very bad at identifying classical signs of deception, we are good at tracking potential motivations to deceive (see e.g. Sperber et al. 2010); thanks to XXX for bringing this to my attention and for very useful comments on several drafts of this paper). Of course, one question that arises at this stage is whether this is actually an actively filtering mechanism or rather just an anti-defeat sensibility. See below for an account that makes sense of the latter.

By the same token, it's plausible that testimonial entitlement requires independent inductive reason to trust the speakers here.

But what about these cases:

**Case 3.** A teacher tells his pupil that two plus two is four.

**Case 4.** A mother tells her child that they are moving to Norway.

Here it is far from clear that taking the speakers' word at face value will amount to gullibility. On the contrary, Greco argues, it is intuitively plausible that, in these cases, the hearers will acquire testimonial entitlement. However, at least on the face of it, neither of the two hearers seems to have access to further reasons to trust the speakers.

John Greco's MAR is motivated by these data. The key idea of his modest form of anti-reductionism is that while testimonial entitlement will sometimes require us to have independent inductive reasons to trust the speaker, at other times, we can have it simply by taking the speaker's word at face value.

Of course, a key question for Greco is exactly when independent inductive reasons are required and when they aren't. It is here that Greco's specifically social answer to SP comes in. More specifically, he claims that whether or not independent inductive reasons are required for testimonial entitlement depends on the social roles of the participants to the conversation. If they belong to the same community of knowers, testimonial entitlement is easy to come by: all they need to do is take the speaker's word at face value. On the other hand, if they do not belong to the same community of knowers, hearers shoulder a more substantive epistemic burden: they need positive reasons to trust the speaker (2015, 292).

According to Greco, participants to testimonial exchanges within a community of knowers engage in a different kind of activity than participants to testimonial exchanges who do not belong to a community of knowers. More specifically, the former centrally involves the *distribution* of information within a community of knowers, whereas the latter centrally involves the *acquisition* of information for the community. Crucially, these two activities have different functions, which, in turn, give rise to different normative requirements. Agents who are engaged in acquiring information for their community have a *gatekeeping function*. They are in charge with letting only genuine information into the system and sifting out misinformation. Note that there is a premium on *avoidance of error* here. That's why we find demanding normative requirements for testimonial entitlement in this kind of case: in order to ensure avoidance of error, positive reasons for trusting the speaker are required. In contrast, agents who are engaged in distributing information within a community are in charge with efficiently distributing high quality

information within a community of knowers. Note that what matters here is *productivity* in the distribution of truths. That's why the normative requirements for testimonial entitlement are laxer here: allowing hearers to take speakers' words at face value is a highly productive way of distributing information within the community.

Unsurprisingly, on Greco's view, the detective and the salesman cases are cases of information acquisition: the corresponding social roles place the agents in importantly different epistemic communities. As a result, testimonial entitlement will be subject to demanding normative requirements: independent inductive reasons for trusting the speaker are needed. In contrast, children and their parents are a paradigmatic case of epistemic agents belonging to the same epistemic community, as are pupils and their teachers. For that reason, the normative requirements at issue in these exchanges are the more lenient ones: taking the speaker's word at face value is just fine, positive reasons are not needed.

The problem for Greco's MAR is that it is not clear that the data do the needed work of motivating his view. That is because, at least at first glance, a classical, strong variety of anti-reductionism, given that it comes equipped with an anti-defeat condition, will do just fine - or, as I am going to try to argue, even better - in accounting for the reductionist intuitions in the Greco cases.

To see this, note that it is plausible that, in both Cases 1 and 2, the hearers have fairly serious undercutting defeaters for believing what they are told: after all, they know that they are talking to people strongly interested in concealing the truth. As such, it will come as no surprise that in order to acquire testimonial entitlement, our agents need positive reason for thinking that these undercutting defeaters do not obtain (in other words, they need defeater defeaters).

The important point here is, of course, that even if SAR grants a *prima facie* entitlement to take a speaker's word at face value, this is compatible with SAR requiring positive reason for trusting the speaker in individual cases. In fact, there is reason to believe that SAR's treatment of Cases 1 and 2 promises to be even better than many reductionist alternatives. After all, what many reductionist accounts of testimonial entitlement require in terms of independent inductive reason for trusting the speaker is something like reason to think that she is in general a reliable testifier (perhaps on the topic at hand). The trouble is that this won't do the trick in all cases. To see this, consider Case 1. It may be that the speaker is in general a reliable testifier (perhaps even on the topic of murder cases). But that's not enough for the agent to acquire testimonial entitlement here. After all, the fact that, as a suspect, she has such excellent reason to lie on this particular occasion will constitute an undercutting defeater despite her general reliability (on the topic). What this suggests is, of course, that we not only need positive reasons, we need a particular variety thereof, the kind of considerations that are able

to defeat the present defeaters. In other words, an explanation in terms of the anti-defeat condition is the most plausible option here.<sup>9</sup>

As a result, there is reason to believe that Greco's MAR fails to improve on SAR on this count. To the contrary: going weak seems counterproductive.

### 3. Burge's A Priori Derivation

In the previous section, I have argued that the motivations put forth by champions of Modest Anti-reductionisms for going weak don't stand up to closer scrutiny. If that is the case, what we are left with is SAR; as such, we had better have a good Strong Anti-Reductionist answer to the Source Problem.

According to Tyler Burge (1993, 2011, 2013), there is reason for optimism: the answer to SP can be identified on purely *a priori* grounds. In a nutshell, here is the view: rationality has the function of generating true content. According to Burge, functional traits have a particular function only if they fulfil it reliably in normal conditions; take the heart, for instance. Were it not to reliably pump blood in your circulatory system in normal conditions, it would likely not have the corresponding function. Since reason has the function of generating true content, it follows that it will do so reliably in normal conditions.

Note that it seems plausible that intelligible propositional expressions indicate a source endowed with rational abilities. If you hear an intelligible message coming your way, it is reasonable to believe that it was a rational source that produced it, rather than, say, that it was randomly generated by some plants in your garden when hit by a gust of wind. You don't need to first gather further positive reasons to attribute rationality to the generating source: that is the default, although defeasible assumption. Since intelligible propositional expressions presuppose rational abilities, intelligible presentations-as-true come *prima facie* backed by a rational source. Now recall that we have seen that, in virtue of having the function of generating true content, rational sources will reliably do so in normal conditions. If that is the case, and if intelligible propositional presentations-as-true indicate a rational source, it follows that intelligible presentations-as-true indicate a *prima facie* source of truth. In virtue of this, according to Burge, we are *a priori prima facie* entitled to take intelligible affirmation at face value (Burge 1993, 472).

---

<sup>9</sup> Note, too, that any variety of reductionism that attempts to vindicate this thought would likely turn out to be implausibly strong; basically, the view would have to demand that hearers have positive reasons to believe every particular speaker is telling the truth on every particular matter of fact. This would eliminate many of the testimonial sources we usually rely upon from the entitlement game.

Note that Burge's view correlates a very ambitious answer to the source problem with a particularly strong variety of anti-reductionism, i.e. with placing no burden on the hearer's shoulders; according to him, the following holds:

SAR: A person is entitled to accept as true something that is presented as true and that is intelligible to him unless there are stronger reasons not to do so" (Burge 1993, 467).

Absent defeaters, then, according to SAR, if something is presented to one as true, one *prima facie* entitled to believe it.

Here is my reconstruction of Burge's argument:<sup>10</sup>

- (1) One is *prima facie* entitled to believe based on intelligible presentations-as-true if they reliably indicate true content in normal conditions.
- (2) Intelligible propositional expressions indicate generation by a source endowed with rational abilities.
- (3) Intelligible presentations-as-true are intelligible propositional expressions.
- (4) Intelligible presentations-as-true indicate generation by a source endowed with rational abilities (from 2 and 3).
- (5) Rational abilities have the function of generating true contents.
- (6) If something X has the function of phi-ing, then X reliably phi-s in normal conditions.
- (7) Rational abilities reliably generate true content in normal conditions (from 5 and 6)
- (8) Intelligible presentations-as-true indicate generation by a source that reliably generates true content in normal conditions (from 4 and 6).
- (9) One is *prima facie* entitled to believe based on intelligible presentations-as-true (from 1 and 8).

I submit that (6) is in need of a closer look. In particular, what I will argue next is that Burge's derivation has difficulties in answering the Source Problem in virtue of lacking support for (6).

There are two problems with (6) that I would like to point out. The first worry concerns reason's dual function: theoretical and practical. It looks as though the latter is, if not reason's primary function, at least on a par with its theoretical function. If that is the case, however, it is not clear why we should believe that, in cases of conflict, the theoretical function of rationality – that of generating true content – will take primacy. After all, the following picture seems like a plausible one: in

---

<sup>10</sup> If I am anywhere close to Burge's line of thought at all, I am highly indebted to Peter Graham's great work on the topic, as well as to his patience and extraordinary clarity in conversation.

most cases, the theoretical function of rationality complements the practical one: true beliefs are good for us, they enable us to find food and avoid predators. On the occasions when that is not the case, the practical function takes precedence over the theoretical one, for reasons pertaining to species preservation. This is just another way to phrase the Source Problem: given reason's dual function, theoretical and practical, why should we believe the former will take precedence over the latter?

Burge is well aware of this problem; his answer comes in two steps. First, he argues that the function of reason of generating truth is transpersonal, that is, independent of the individual's personal interests. Believing against the evidence, for instance, even when it is all things-considered rational to do so, involves a failure of rationality — of theoretical reason. Here is Burge:

One of reason's primary functions is that of presenting truth, independently of special personal interests. Lying is sometimes rational in the sense that it is in the liar's best interests. But lying occasions a disunity among functions of reason. It conflicts with one's reason's transpersonal function of presenting the truth, independently of special personal interests (2013, 242-243).

Second, according to Burge, when a trait has more than one function, normal conditions will be such that the relevant functions will not conflict with each other. In normal conditions, reason will generate both true content and prudential benefits reliably: one will not have to lie in order to promote prudential goods. As such, in normal conditions, rational sources will generate true content:

I think that generic rationality has practical and impersonally theoretical dimensions. I think that, *prima facie*, when a speaker fails to tell the truth because of special interests, the speaker crosses rationality in one significant dimension. Just as a generically rational individual can be irrational on particular occasions, a generically rational individual can be all things considered rational in the particular case, while rationally lying. But the impersonal function of rationality is compromised. So there is some compromise on generic rationality (2013, 266).

The claim is that, when a trait has more than one function, in normal conditions, they will not conflict, and thus, in principle, they will all be reliably fulfilled.

Now, the worry is that this does not seem to hold for all functional traits: conflicting functions are hardly a rarity in nature. For a straightforward exception to this, consider, for instance human sexual

organs, and their functions of excretion and reproduction. Or take the elbow in carnivorans, for instance: cursorial locomotion and manual manipulation are conflicting functions; the price paid by pursuit predators for limbs adapted for speed and locomotor efficiency is the loss of the ability to supinate the forearm, which is essential for grappling with prey. If conflicting functions are not an exception in nature, however, it seems highly unlikely that the Burge derivation will go through on a priori grounds: it looks as though, when a trait has two functions, it need not be the case that, in normal conditions, the two do not come in conflict. To the contrary, conflict might be the default state.

The second worry with (6) concerns the universality of the reliability assumption: (Graham Forthcoming) points out that it is not clear that just because a trait T has a function F, it follows that it will reliably fulfil it in normal conditions. Take sperm, for instance: only very few among thousands of sperms actually succeed in fertilizing the egg. Similarly, think of the camouflage mechanism against predators present in many animals, or of danger-detection mechanisms in rabbits. On most occasions, function will not be fulfilled: most animals will not survive predation, and most of the time mechanisms in rabbits will fail to detect real danger. Still, importantly, these traits, while not strictly speaking reliable, they are effective: they are reliable enough.

If that is the case, however, it looks as though one cannot establish the reliability of reason in fulfilling its function of generating true content a priori: rather, one needs to investigate empirically whether reason belongs with hearts, or rather with sperm in this respect.

One thing that one could do in order to save the Burge derivation from the Graham worry would be to check whether the view survives if we substitute *reliable enough* in the relevant premise. If reason generates true content *reliably enough*, would that not suffice for epistemic entitlement? Here is the revised Burge argument:

- (1') One is prima facie entitled to believe based on intelligible presentations-as-true if they *reliably enough* indicate true content in normal conditions.
- (2) Intelligible propositional expressions indicate generation by a source endowed with rational abilities.
- (3) Intelligible presentations-as-true are intelligible propositional expressions.
- (4) Intelligible presentations-as-true indicate generation by a source endowed with rational abilities (from 2 and 3).
- (5) Rational abilities have the function of generating true contents.
- (6') If something X has the function of phi-ing, then X *reliably enough* phi-s in normal conditions.
- (7') Rational abilities *reliably enough* generate true content in normal conditions (from 5 and 6')

- (8') Intelligible presentations-as-true indicate generation by a source that *reliably enough* generates true content in normal conditions (from 4 and 6).
- (9) One is prima facie entitled to believe based on intelligible presentations-as-true (from 1' and 8').

Think of sperm: what does reliable enough stand for? Well, it looks as though what is at stake will be something like 'enough for reproduction purposes'. So it looks like the 'enough' at stake will be domain specific: enough for the relevant type of effectiveness, in the relevant domain. Couldn't we, then, have a similar picture for rationality, whereby it generates true content reliably enough for our epistemic purposes? And would that not be tantamount to saying that it does so reliably enough for epistemic entitlement?

Unfortunately, things are more complicated than this. And here is why: it looks as though the 'enough' at stake in 'reliable enough' will likely be a biological, prudential one rather than an epistemic one. Why is it, after all, that a trait needs to fulfil its function reliably enough to begin with? Well, one obvious answer is because that contributes to the trait's continuous existence: the heart, in order to persist, needs to pump blood with near maximum reliability. In contrast, sperm need not do more than succeed once in a while, enough to insure continuity of species. If hearts would fail to be reliable enough from a biological, prudential point of view, they will likely be discontinued. Similarly with sperm. But if that is the case, plausibly, in the case of rationality too, the relevant threshold setter will be biological, prudential rather than epistemic: our rational capacities will generate enough true beliefs to keep us alive and well. Of course, the two may coincide. But then again, they might not. It might be that, for mere survival, much less in the way of ratio of true to false beliefs is needed than for epistemic entitlement. True beliefs about the whereabouts of food and approaching predators need figure in the relevant set. Arguably, true beliefs about the year of the fall of Constantinople, or about the highest peak in the Carpathians need not. In any case, the point is that it is not clear – and even less plausibly available on a priori grounds – that the derivation gives us the relevant, epistemic 'enough' when it comes to true beliefs generated by our rational capacities.

If that is the case, though, when it comes to epistemic entitlement, Burge misses support for (1'): if the 'enough' at stake need not be epistemic, one will not be prima facie *epistemically* entitled to believe based on intelligible presentations-as-true if they *reliably enough* indicate true content in normal conditions. Rather, the entitlement at stake will be non-epistemic.

To sum up: we have seen that, just because reason has the



function of generating true content, it need not follow that (1) it fulfils it reliably enough for epistemic purposes, and (2) that reasons' prudential function does not interfere with the reliability of the theoretical function. If that is the case, though, it looks as though the solution to the Source Problem does not afford Burge's envisaged a priori derivation.

#### **4. Three Ways to Read the Source Problem**

Now, let's have a closer look at the Source Problem. Recall, the thought was that people's freedom to lie, in conjunction with the rationality of their doing so to further their own best interest was taken to generate difficulties for anti-reductionism. In what way, though? While intuitive enough, SP remains vague about this. Too vague, that is: after all, as Goldberg well puts it, other epistemic sources are fallible too, and this does not rob them of entitlement-generating force.

What I will do next is go through three possible ways to understand SP, in order to better figure out what the task for the anti-reductionist consists in.

##### *3.1 The Reliability Reading*

One very straightforward way to read SP is as concerning the reliability of testimony. Roughly, the thought would go as follows: epistemic entitlement requires reliable formation. However, people are free to lie; also, they will do so whenever this furthers their interests. Since the interests of the hearers very rarely align with the interests of the speakers, testimony is not reliable enough to generate epistemic entitlement.

Importantly, though, as straightforward as this reading might be, it cannot be the correct one: SP does not concern the reliability of testimony. The question is not: people lie a lot, therefore why trust them? To see this, note that, after all, testimonial knowledge is extremely ubiquitous, which seems to suggest that testimony is a very reliable source: it is fairly widely accepted that reliable formation is a necessary condition on knowledgeable belief.

Second, note that the reliability (or lack thereof) of testimony is an empirical matter: we can go out in the world and measure it. But we would surely not want to say that the beef between reductionists and anti-reductionists in the epistemology of testimony is an entirely empirical issue; that is, it is surely not the case that all we need in order to adjudicate between these two views is some statistics.

##### *3.2 Trust vs. Reliance*

Here is another way to look at the source problem. Recall that what the reductionist is asking for is reasons to *trust* the testifier. Maybe the notion of trust itself is one that, independently of how reliable the speaker might be, or of whether she is interested in telling the truth or not, requires epistemic work on the part of the hearer. Anette Baier's (1986, 235) classical work on the issue might help here.<sup>11</sup> According to Baier, there is an important distinction between trust and mere reliance: the former requires a leap of faith, and is, as such, a more emotionally charged attitude. "trusting can be betrayed, or at least let down, and not just disappointed" (1986, 235). One can rely on inanimate objects, such as alarm clocks; but when they break, one is not betrayed, although one may be disappointed.

Accordingly, maybe one way to frame the source problem is as follows: just as trust and reliance can come apart, so can the corresponding conditions for entitlement. In particular, one could think entitlement to trust, in contrast to entitlement to rely on, requires positive reasons. Testimony is an interpersonal affair; speakers, as opposed to alarm clocks, have the freedom to lie if this furthers their own interest. This, in turn, suggests that one needs to take a leap of faith in testimonial exchanges that is not at stake on relying on instruments. What is at stake, then, in testimony, is entitlement to *trust* the speaker; as such, the thought goes, positive reasons to this effect are needed. After all, why would I invest a complete stranger with the kind of trust that I only bestow upon my closest and dearest without having any good reason to do so?

Even if we accept this demanding picture for entitlement to trust, however, there is still a long way to the Source Problem. To see why, first, we need to distinguish between trust as a two-place relation – A trusts B – and three-place trust – A trusts B to C. Note that two-place trust is a more demanding affair: I can, of course, trust you to do the dishes, although I don't generally trust you. Relatedly, three-place trust is a less emotionally charged attitude; it is one thing to trust Mary – which might require some happy history between us – and quite another to trust Mary to do the dishes.

Note, too, that three-place trust is still importantly different from mere reliance: I can rely on you to do the dishes – say, because I have no choice – but, at the same time, not trust you to do so.

Furthermore, the attitude at stake in Baier's picture is what Paul Faulkner (2011) calls affective – emotionally charged – trust, as opposed to mere predictive trust. However, the only interest in trust that reductionists have is trust in its predictive dimension: nothing more emotionally sophisticated than that seems to be needed. The latter, however, seems to require much less in the way of support. To see this,

---

<sup>11</sup> See, also, (Faulkner 2017).

note that predictive trust is something that, in one way or another, we manifest all the time in our everyday life, maybe without even giving it much thought: I trust my baker to give me a loaf of bread in return for my 2 pounds, although I know nothing about him. I trust the drivers to stop the cars at the red light, although I have no idea who's driving. I trust the city hall to not have replaced the real oak tree in front of my office window with an undistinguishable fake one. I trust the nanny to take the kids to school, I trust the electrician to repair the circuit. And so on.<sup>12</sup> Predictive trust is everywhere, and no (accessible) positive reasons seem to be required to be entitled to it. To see this, note that, after all, small children do not have the cognitive sophistication to reflect upon the trustworthiness of the drivers, bakers etc, and still, they seem perfectly entitled to place three-place trust in all these people.

### *3.3 SP as a Problem of Cooperation*

Paul Faulkner (2011) thinks the Source Problem is essentially a 'problem of cooperation', a variety of the Prisoner's Dilemma for testimonial entitlement (2011: 6):

[S]peakers and audiences have different interests in communication[...] Our interest, qua audience, is learning the truth. Engaging in conversations as to the facts is to our advantage as speakers because it is a means of influencing others: through an audience's acceptance of what we say, we can get an audience to think, feel, and act in specific ways. So our interest, qua speaker, is being believed...because we have a more basic interest in influencing others....[T]he commitment to telling the truth would not be best for the speaker. The best outcome for a speaker would be to receive an audience's trust and yet have the liberty to tell the truth or not. (2011: 5-6)

According to Faulkner, then, the default position for speakers bears no commitment to telling the truth. If that is the case, the default position for hearers bears no entitlement to believe. Therefore, SP.

It is important to understand the significance of the '*default*' position at stake in the argument. Recall that what is at stake in the debate between Reductionism and Anti-reductionism is *prima facie* entitlement. Both views will allow, for instance, that if there are defeaters present at the relevant context, hearers are not entitled to believe. This explains the concern with what the default, or starting position is in

---

<sup>12</sup> These statements can all be recast as belief statements: I believe that the baker will give me a loaf of bread in return for my 2 pounds etc. As such, arguably, it is not clear that the attitude at stake is even trust proper (Paul Faulkner, p.c.).

testimonial exchanges: all else absent, are we entitled to believe what we are being told? Anti-Reductionism says ‘yes’; Reductionism says no, you need to go out in the world and look for further support for your belief.

Faulkner’s view aims to explain the latter position: all else absent, the default for the speaker is not telling the truth. Therefore, the default for the speaker is not entitlement to believe.

I agree with Faulkner that this is the most plausible way to read SP. In what follows, though, I will attempt to reconstruct Faulkner’s argument, in order to point out one important feature of this reading of the Source Problem. Here is one, particularly strong reading of it:

- (1) Hearers are interested in truth; speakers are interested in being believed.
- (2) The *default* position for speakers is seeing to their own interests rather than to the interests of the hearers.
- (3) Therefore, it is not the case that the default position for speakers is telling the truth (from 1 and 2).
- (4) The default position for hearers is trust only if the default position for speakers is telling the truth.
- (5) Therefore it is not the case that the default position for hearers is trust (from 3 and 4).

Here is the main worry I have with this reading of the Faulkner view: if the reconstruction above is correct, the conclusion fails to follow. In particular, the problem is with premise (3), which is not supported by (1) and (2). That is because being interested in being believed does not exclude also being interested in telling the truth. Speakers might still – by default - also be interested in telling the truth, on independent grounds, that is, independently of their concern (or, rather, lack thereof) with hearer’s interests. If that is the case, telling the truth will be default for hearers, therefore trust will be default for hearers. In other words, hearers do not need the speakers to provide the goods from altruistic motivations; they just need them to provide the goods.

Of course, the talk in the argument is of a default position. And if you have a reason to believe that a speaker’s interest in the communicative exchange is getting you to believe what they say *and you have no other information about the speaker*, that this reason is in itself a reason to not believe what is said.<sup>13</sup> Note, though, that this amounts to a weaker reading of Faulkner’s Problem of Cooperation, whereby it is a *challenge* to Anti Reductionism rather than a problem for the view; a challenge to explain why the speakers would, *by default*, tell the truth rather than not, given that they are rational, self-interested agents, and independently of their lack of concern for the interests of the hearer. If a good answer to

---

<sup>13</sup> Thanks to xxx for pressing me on this.

this question can be found, we will get the game theoretic picture that we wanted to begin with: the default position for hearers will be entitlement to believe, in virtue of the default position for hearers being truth telling.<sup>14</sup>

The next section answers Faulkner's challenge. In a nutshell, the thought I aim to develop in the rest of the paper is a fairly simple one: the existence of the social contract sets the default position for agents to social norm compliance – i.e., all else absent, it is in their best interest to comply. As such, the default position for receivers of their services within the social contract is entitlement to trust that they will comply. We are entitled to trust the baker and the driver to respect the social norms in virtue of the fact that the social contract they are acting in makes it such that it is in their best interest to do so; similarly, I will argue, we are entitled to believe what we are being told in virtue of the fact that the default position for speakers – within the social contract - is epistemic norm compliance.

To do that, in what follows, I will propose a novel, social strong anti-reductionist account, which I dub *Testimonial Contractarianism*. The account put forth is strongly anti-reductionist, in that it argues that speakers are prima facie entitled to believe what they are being told. It is a social anti-reductionism, in that, contra Burge, it takes the source of such entitlement to be a posteriori, sourced in the social contract.<sup>15</sup>

## 5. Testimonial Contractarianism

### 4.1 *Enlightened Self-Interest*

---

<sup>14</sup> Faulkner (2011) himself offers a reductionist (in the sense restricted employed by this paper, i.e. non-credulist) variety of a social contract – based account. In a nutshell, on Faulkner's view (2011, chapter 7), the problem of cooperation is solved by our intrinsically valuing trust, which is the result of social normativity. On this view, we see another needing information as in itself a reason to give them the information they need. This norm sets up our concept and practice of trust. And it is this practice that ensures that our trusting others means we have a reason to believe them. Trust supplies a reason — and so satisfies the reductive demand for a reason — only because and insofar as there is a practice of trust.

Since this paper's focus is restricted to anti-reductionist responses to SP, I will not engage with Faulkner's view here. Further research will focus on comparing reductionist to anti-reductionist social normativity-based accounts. See, however, Graham (2012) for discussion.

<sup>15</sup> Several contractarians (e.g. Gauthier 1987) attempt to argue for a purely rational, a priori basis of the social contract itself. I will not take a stance on this issue here. Importantly, though, if they are right, this paper's answer to the Source Problem is a priori, in virtue of its being grounded in the social contract which, in turn, on this picture, would be grounded in a priori norms of rationality. I must confess that I am rather pessimistic that this picture will go through, however: it seems to me as though it is only rational to enter the social contract given the kind of vulnerable creatures we are, which seems a contingent matter of fact. Thanks to XXX for illuminating discussions on this issue and many others in this paper.

Socrates is first to make contractarian remarks in *Crito*: according to him, members within a society implicitly agree to the terms of a kind of social contract by their choice to stay within the society. Social and political norms, then, receive their legitimacy from this implicit ‘signing’ of the social contract by all the members of society.

Contractarianism is a normative view that draws its resources from the social contract. Thomas Hobbes (1651) is the most notable historical figure explicitly defending such a view. In a nutshell, the Hobbesian story goes as follows: in a (actual or hypothetical) primitive unstructured social order, what he dubs a ‘state of nature’, individuals have unlimited natural freedoms to seek and further their self-interest. However, this also includes the freedom to harm all who threaten one's own self-preservation. This, in turn, likely leads to a general conflictual, unsafe environment, unfriendly to life preservation and nurture. It is therefore, he argues, in an individual's rational self-interest to voluntarily give up some of his freedom of action in order to obtain the benefits provided by the formation of social structures and civil rights. In a nutshell, the social contract binds individual liberties, with an aim to an increased security of the participants in it.

One can distinguish between two contractarian claims: a strong and a weak one. The strong claim takes contractarian explanations to account for the very legitimacy of moral, political or social norms. I will put this claim aside here. The (weak) contractarian claim that I am interested in merely concerns motivational aspects: the social contract explains the fact that free, self interested agents comply with extant moral, political or social norms, whichever their source, rather than defect when it suits them.

Let us dub this ‘Compliance Contractarianism’. Again, according to this view, the social contract explains why people comply with norms when it looks as though doing so goes against their immediate interest. This view says nothing about the source of the relevant normativity, and certainly does not take it to reduce to social normativity.

Now, the reason why Compliance Contractarianism is particularly interesting here, given the concerns of this paper, is the fact that it historically concerns itself with answering the problem of cooperation: note that, for each and every agent within the social contract, it looks as though it is, as a matter of fact, in her best interest to defect. For our agent, there are two possible scenarios: either everyone else complies with the norms, or they don't. If they don't, she's better off not complying either. If they do, it is best for her if she defects. In short, though rational individuals might consider it in their best interest to enter the contract in order to ensure a great level of security and welfare for themselves, once inside, they will find that complying with the norms might not be the most profitable way of handling things.

And still, just like with testimonial knowledge, we can see with

our own eyes that the social contract<sup>16</sup> does, in fact, in some way, overcome this problem; Amartya Sen's famous "stranger and the letter" example nicely illustrates this.<sup>17</sup> Suppose you are headed for the post office to deliver a letter and a stranger approaches you and asks for directions to get to the train station. Your immediate self-interest should then guide you to lie to the poor man that the train station is right by the post office, and ask him, while he is heading towards it, to also deliver your letter. This will certainly save you a trip. Though, I suppose everybody would agree, no one would do such a thing, but they would rather give good directions to the stranger and deliver the letter themselves.

The contractarian literature features a fairly widely endorsed explanation for this datum, first defended by David Gauthier (1987). In a nutshell, according to Gauthier, what does the trick is enlightened self-interest: he argues that rational individuals will choose to dispose themselves to be constrained self-interest maximizers rather than straightforward self-interest maximizers. When rational, people will choose to disregard their immediate gain for future, more substantive gain. To see the plausibility of this claim, take some straightforward examples of enlightened self-interest maximization our society is involved in on an everyday basis: spending one's youth working through long years of education, rather than lying on the beach playing Tetris; going to work everyday, rather than staying at home watching a season of one's favorite TV show; buying a house rather than taking a luxury vacation in Hawaii; and so on.<sup>18</sup>

Further confirmation notably comes from game theory. According to people like Gary Bolton (1991), Jack Ochs and Alvin Roth (1989), the existence of the social norm affects the utility profile of a particular context. There is future gain to be had from conforming with the social norm – in terms of good reputation, social approval, decreasing risk of being subject to sanctioning. As such, all else absent,

---

<sup>16</sup> Here and below the social contract refers to society in general, not to a particular country or other such explicitly delineated social group. I take it to be intuitive that non-members of our society – such as Lackey's Alien – will fail to be motivated by social norms, and that we lack entitlement to believe they will comply with them.

<sup>17</sup> See McQuaig (2001) for discussion.

<sup>18</sup> Faulkner (2017) worries that Gauthier's solution hinges on a belief about the trustee's motivational dispositions, namely that the trustee is a constrained maximizer, not a straightforward one, *and recognised as such*. If so, in a situation of ignorance, the rational thing to do, according to Faulkner, would be to presume that the other is a straightforward maximizer and avoid relying on them. Not so, argues Gauthier. The inability to detect straightforward maximizers would not matter if the proportion of constrained maximizers in the population were significantly large. Of course, even if this is so, Faulkner argues, it will be an empirical matter of fact; as such, Gauthier's argument will fail to go through on a priori grounds. As announce earlier in this paper, however (fn. 13), the account defended here does not rest on the a priori nature of Gauthier's derivation, and thus will not be affected by this problem.

in social contexts, norm conformity is the default position, it enjoys default rationality.

Furthermore, leaving the default position requires fairly serious incentive; the stakes need to be rather high for the benefits of defecting to outweigh the benefits of conforming. That is, the expected payoff needs to be as high as to make defecting rational, when combined with the probability of being sanctioned. To see this, consider Ultimatum games:

**ULTIMATUM:** There are two players, a Proposer and a Responder. The Proposer is given a sum of money – say, one hundred pounds. He then must propose a split of the money between the Proposer and the Responder. The Responder’s job is to accept or refuse the split. If accepted, both parties receive the amount proposed. If refused, no one gets anything. As such, both parties are better off if the Responder accepts the split.

Note that, as the case is described, absent any further considerations, the Proposer should propose one pound for the Responder and 99 pounds for the Proposer. Also, the Responder should accept. After all, the thought goes, one pound is better than nothing.

Across a very wide variety of human cultures, however, that is not what happens. Instead, the Proposer tends to offer something in the vicinity of a 40/60 split. Furthermore, in cases where the Proposer does offer a much smaller split to the Responder, the Responder tends to refuse.

The explanation of his behavior – which seemingly violates rational choice axioms – is that the presence of the social norm changes the utility profile of the case, such that more in the way of expected payoff is required to outweigh the utility contribution of norm conformity. Studies show that, in higher stakes cases, unsurprisingly, Respondents’s behavior changes dramatically: “ ... among respondents we find a considerable effect of stakes: while at low stakes we observe rejections in the range of the extant literature, in the highest stakes condition we observe only a single rejection out of 24 responders” (Andersen et al. 2011).

Absent substantial payoff for defecting, however, the default rational position within the social contract is social norm compliance. This, of course, goes a long way in explaining the intuitive easy entitlement to trust that norm conformity will obtain in social settings, such as the baker’s, or traffic, or the like. The presence of the social norm will modify the utility profile such that, absent strong incentives to defect, by default, the baker is rational to give me the loaf of bread in exchange of my two pounds, and the drivers are rational to stop at the red light. Conversely, I am rational to trust that they will.



#### 4.2 Testimonial Contractarianism

By now, it should come as no surprise that I want to propose that a similar picture holds for testimonial exchanges too. Note that it is fairly<sup>19</sup> uncontroversial that the speech act of assertion is governed by an epistemic norm. Note, also, that some epistemic norms are also social norms; plausibly, norms governing speech acts are among the main candidates for this double status. That is not to say that the relevant epistemic normativity reduces to social normativity; rather, the claim is a coincidence claim: some norms are, at the same time, epistemic and social. Plausibly, the norm of assertion will figure among such norms: it is, after all, not only epistemically unacceptable to lie, for instance, but also socially unacceptable.

Several people think knowledge is the norm of assertion: one should only assert that *p* if one knows that *p* (KNA).<sup>20</sup> Others disagree and impose weaker conditions (justification, truth), or stronger conditions (certainty), or even contextually variant conditions on permissible assertion.<sup>21</sup>

Whatever their preferred norm might be, however, most people in the debate agree upon one minimal, happy Gricean thought: since it should be able to account for assertion's epistemic aim, or function, it features a positive epistemic status. In other words, the norm is not 'Assert that *p* only if *p* is false!', nor 'Assert that *p* only if you have no reason to think *p* is true!' That this much hearer friendliness exists is accepted by pretty much everyone in the debate. For this reason, the argument to follow can go through by employing any one of the accounts of the normativity of assertion proposed in the literature. For the sake of simplicity, but also because it is the account I favour, in what follows, I will go with the knowledge norm (KNA). According to the latter, one must: assert that *p* only if one knows that *p*.

Here is what I will argue next, in a nutshell: given that KNA is also a social norm, all else absent, in virtue of the social contract, conforming to it enjoys default rationality for the speakers. If conforming to KNA enjoys default rationality for the speakers, hearers are by default entitled to believe what they are being told.

---

<sup>19</sup> With the notable exception of Cappelen (2011).

<sup>20</sup> Champions of KNA include Peter Unger (1975), Michael Slote (1979), Keith DeRose (2002), John Hawthorne (2004) and most famously Timothy Williamson (2000). I have also defended KNA in several places (e.g. Author 2016a, Author 2016b, Author 2016c, Author 2017(a), Author Forthcoming). For a reductionist view of testimony centred around KNA see (Fricker Forthcomingb).

<sup>21</sup> See Douven (2006) and Lackey (2008) for a justification norm of assertion. For the truth norm of assertion, see e.g. Weiner (2005) and Whiting (2013). Stanley (2008) defends a certainty norm of assertion. See e.g. Brown (2010), Gerken (2012), Goldberg (2015) and Greenough (2011) for context sensitive norms.

To lay my cards right on the table, here is the strong social anti-reductionist view defended in the paper:

**Testimonial Contractarianism:** In virtue of the social contract, hearers are prima facie entitled to form beliefs based on speaker's assertions.

The view is a Burge-style strongly anti-reductionist view, in that it imposes no epistemic burden on the hearer: absent defeaters, she is entitled to believe on mere say-so. It is a Goldberg, Graham and Greco style anti-reductionism in that it takes the source of this entitlement to be social.

Testimonial exchanges are distinctively social phenomena; by default,<sup>22</sup> they happen within and as part of the social contract. The fact that a social norm requiring knowledge for permissible assertion is in place, in virtue of the social contract being in place, changes the utility profile of testimonial exchanges. In virtue of its also being a social norm, by contractarian lights, conforming to KNA will enjoy default rationality on the side of the speaker. All else absent, speakers have incentive to assert knowledgeably. Just like stopping at the red light is the default for drivers, and giving me a loaf of bread in exchange for my two pounds is the default for the baker, in virtue of the social contract being in place, the default position for speakers is asserting knowledgeably.

If that is the case, the default position for hearers is entitlement to believe, just like the default position for pedestrians is to cross the street on a green light, and the default position for me is to trust that my baker will give me the loaf of bread in exchange of my two pounds.

Of course, just like with any other social norms, strong reasons can come and override the requirements of KNA. If it is urgent to take your wife to the hospital, you might break the traffic norms. Similarly, if he desperately needs both the bread and my money for his children, the baker might defect and not respond in the expected way. Similarly, if it serves your interests to lie to me, you will. Note, though, crucially, that in

---

<sup>22</sup> One natural worry at this point is that I am building too much into this 'default' position. After all, wasn't this supposed to be something pertaining to prima facie entitlement, in an empty utility room, as it were? The answer is 'no'. Some information about utilities needs to be allowed into the default position in order to get the Source Problem off the ground to begin with. After all, the fact that speakers are interested in being believed rather than spreading truths, for instance, is a distinctively social phenomenon, pertaining to how our societies happen to function – i.e., in this case, pertaining to the fact that influencing others tends to be to one's one advantage. Building this into the default position presupposes the existence of the social contract in play, at least in the weak variety thereof which is at stake in this paper. Many thanks to xxx for pressing me on this one.

Note, also, that without such assumptions in play, the Source Problem vanishes. After all, what we get in an empty utility room is merely a speaker trying to waist as little utility for no gain whatsoever. It is easily arguable that in such a scenario, the rational choice is to remain silent or, alternatively to tell the truth, just in virtue of the fact that it's easier, it takes upless cognitive resources.

all the above cases, the live possibility of norm violation does not affect my default entitlement to trust in norm conformity.

Of course, the discussion here concerns merely the default position for hearers, as sourced in the default position for speakers. As such, all we get from this normative story is prima facie entitlement to believe. Defeaters might and will affect this position. If I know your wife needs to get to the hospital urgently, or that you stand to gain from lying to me, I am not entitled to expect norm conformity anymore. You will have exited the all-else-absent default position, in virtue of further incentives overriding the ones sourced in conforming to the relevant norm. Correspondingly, I will not enjoy default entitlement to trust you.

## 6. Conclusion

I have argued that you can trust me to tell you the truth. You don't need to know who I am, whether I have a good track record in these respects, or whether I might have good reason to lie or not. Nor do you need to have reason to believe in the reliability of testimony in general. You don't even have to give all these things any thought. All you need is to comprehend my assertion; next thing you know, you're entitled to believe it. Such are the workings of the social contract.

## References

- Andersen, S., Ertac, S., Gneezy, U., Hoffman, M., & List, J. A. (2011). Stakes Matter in Ultimatum Games. *The American Economic Review*, 3427-3439.
- Author (2016a). Author's Work.
- Author (2016b). Author's Work.
- Author (2016c). Author's Work.
- Author (2017a). Author's Work.
- Author (2017b). Author's Work.
- Author (Forthcoming). Author's Work.
- Bolton, G. (1991). A comparative model of bargaining: theory and evidence. *Levine's Working Paper Archive* 263, David K. Levine.
- Brown, J. (2010). Knowledge and Assertion. *Philosophy and Phenomenological Research* 81: 549-566.
- Burge, T. (1993/2013). Content Preservation. In *The Philosophical Review* 104: 457-488. Reprinted in Burge 2013.
- Burge, T. (2013), Cognition Through Understanding: Self-Knowledge, Interlocution, Reasoning, Reflection. *Philosophical Essays*, Vol. 3, Oxford: Oxford University Press.

- Burge, T. (2011/2013). Postscript: ‘Content Preservation.’ In *Cognition Through Understanding*: 254-284.
- Cappelen, H. (2011). Against Assertion. In Jessica Brown and Herman Cappelen (eds.), *Assertion: New Philosophical Essays*. Oxford: Oxford University Press.
- DeRose, K. (2002). Assertion, Knowledge, and Context. *Philosophical Review* 111: 167-203.
- Douven, I. (2006). Assertion, Knowledge, and Rational Credibility. *Philosophical Review* 115: 449-485.
- Faulkner, P (2011). *Knowledge on Trust*. Oxford: Oxford University Press.
- Faulkner, P. (2017). The Problem of Trust. in *The Philosophy of Trust* eds. Faulkner and Simpson, OUP.
- Fricker, E. (Forthcominga) Unreliable Testimony. In *Alvin Goldman and his Critics*, Hilary Kornblith and Brian McLaughlin (eds.), Blackwell.
- Fricker, E. (Forthcomingb). Inference to the Best Explanation and the Receipt of Testimony. *Best Explanations: New Essays on Inference to the Best Explanation*. Ted Poston and Kevin McCain (eds.): OUP
- Fricker, E. (1995). Critical Notice: Telling and Trusting: Reductionism and Anti-Reductionism in the Epistemology of Testimony. *Mind*, 104: 393–411.
- Gauthier, D (1987) *Morals by Agreement*. Clarendon Press.
- Gerken, M. (2013). Same, Same but Different: The Epistemic Norms of Assertion, Action, and Practical Reasoning. *Philosophical Studies* 168: 725-744.
- Goldberg, S. (2010). *Relying on Others*. Oxford: Oxford University Press. 2010.
- Goldberg, S. (2014). Interpersonal Epistemic Entitlements. *Philosophical Issues*, vol. 24/1: 159–183
- Goldberg, S. (2015). *Assertion: The Philosophical Significance of a Speech Act*. Oxford: Oxford University Press.
- Graham, P. J. (2006). Can Testimony Generate Knowledge? *Philosophica* 78: 105-127.
- Graham, P. (2012). Testimony, Trust, and Social Norms. *Abstracta Special Issues* 6 (3): 92-116
- Graham, P (Forthcoming). Sincerity and the Reliability of Testimony: Burge on the A Priori Basis of Testimonial Entitlement. In Stokke and Michaelson (eds.) *Lying: Language, Knowledge, Ethics, Politics*: Oxford University Press.
- Greenough, P. (2011). Truth-Relativism, Norm-Relativism, and Assertion. In J. Brown and H. Cappelen (eds) *Assertion: New Philosophical Essays*, Oxford: OUP, 2011.
- Hawthorne, J. (2004). *Knowledge and Lotteries*. Oxford: Oxford University Press.
- Hobbes, T. (1985) [1651]. *Leviathan*. C.B Macpherson (Editor). London: Penguin Books (1985)

- Lackey, J. (1999): Testimonial Knowledge and Transmission. *The Philosophical Quarterly* 49: 197, 471-90.
- Lackey, J. (2008). *Learning From Words*. Oxford: Oxford University Press.
- Michaelian, K. (2010). In Defence of Gullibility: The Epistemology of Testimony and the Psychology of Deception Detection. *Synthese* 176(3): 399–427.
- Ochs, J. and Roth, A. E. (1989). An experimental study of sequential bargaining. *American Economic Review*, 79(3): 355–84.
- Sen, Amartya (1970). *Collective Choice and Social Welfare*. Holden-Day.
- Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origg, G., & Wilson, D. (2010). Epistemic vigilance. *Mind & Language*, 25(4), 359-393.
- Slote, M. (1979). Assertion and belief. In Dancy, J., editor, *Papers on Language and Logic*. Keele University Press, Keele.
- Stanley (2008). Knowledge and Certainty. *Philosophical Issues*, 18: 33-55.
- Unger, P. 1975. *Ignorance: A Case for Scepticism*. Clarendon, Oxford.
- Weiner, M. (2005). Must we know what we say? *Philosophical Review*, 114:227– 251.
- Whiting, D. (2013). Stick to the facts: on the norms of assertion. *Erkenntnis*, 78:847– 67.
- Williamson, T. (2000). *Knowledge and its Limits*. Oxford: Oxford University Press.