TESTIMONY, CONTEXT, AND MISCOMMUNICATION
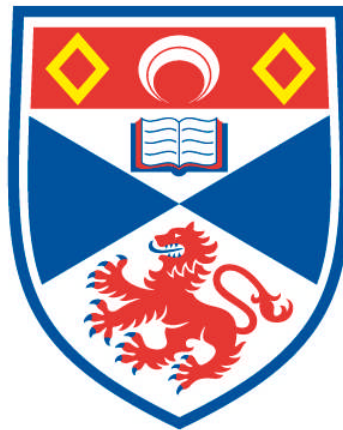
Andrew Peet

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews

2015

Full metadata for this item is available in
Research@StAndrews:FullText
at:
http://research-repository.st-andrews.ac.uk/

Please use this identifier to cite or link to this item:
http://hdl.handle.net/10023/7705

This item is protected by original copyright

Testimony, Context, and Miscommunication.

Andrew Peet

A thesis to be submitted to the
University of St Andrews
for the degree of
Doctor of Philosophy.

Departments of Philosophy

School of Philosophical, Anthropological and Film Studies

University of St Andrews

June 2015.

I, Andrew Peet, hereby certify that this thesis, which is approximately 66000 words in length, has been written by me, and that it is the record of work carried out by me or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for a higher degree.

I was admitted as a research student in September 2011 and as a candidate for the degree of Doctor of Philosophy in September 2011; the higher study for which this is a record was carried out in the University of St Andrews between 2011 and 2015.

Date: Signature of candidate:

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of ……… in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date: Signature of supervisor:

In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that my thesis will be electronically accessible for personal or research use unless exempt by award of an embargo as requested below, and that the library has the right to migrate my thesis into new electronic forms as required to ensure continued access to the thesis. I have obtained any third-party copyright permissions that may be required in order to allow such access and migration, or have requested the appropriate embargo below.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis: No embargo on print copy. No embargo on electronic copy.

Date:

Signature of candidate:

Signature of supervisor:

Abstract

This thesis integrates the epistemology of testimony with work on the epistemology, psychology, and metaphysics of language. Epistemologists of testimony typically ask what conditions must be met for an agent to gain testimonial justification or knowledge that *p* given that *p* has been asserted, and this assertion has been understood. Questions regarding the audience's ability to grasp communicated contents are largely ignored. This is a mistake. Work in the philosophy of language (and related areas) suggests that the determination and recovery of communicated contents is far from straightforward, and can go wrong in many ways. This thesis investigates the epistemology of testimony in light of this work, with a special focus on miscommunication.

The introduction provides a brief overview of some relevant work on testimony, the philosophy of language, and psychology, and argues that there is good reason to investigate the three. One obvious problem in this area is that if testimonial knowledge requires knowledge of what is said then the risk of miscommunication will block testimonial knowledge. Chapter two argues that testimonial knowledge does not require knowledge of what is said. The remaining four chapters discuss problems which do arise from miscommunication. Chapters three and four focus on the epistemic uncertainty of communication with context sensitive terms. Chapter three argues that many beliefs formed on the basis of context sensitive testimony are unsafe and insensitive. Chapter four argues that speakers often have plausible deniability about the contents of their assertions. Chapters five and six explore types of miscommunication which arise as a result of background mental states affecting our linguistic understanding. Chapter five explores the social/ethical consequences of this, arguing that certain groups are disproportionately subject to harmful misinterpretation. Chapter six argues that testimonial anti-reductionists make the wrong predictions about a range of cases of cognitive penetration.

Acknowledgements.

Chapter 1: Introduction.

0. Introduction.

The epistemology of testimony is the study of the beliefs we form through being told something. The study of the epistemology of testimony includes questions such as the following: under what conditions do we gain knowledge through testimony? Under what conditions are we justified in accepting testimony? Is testimony a generative source of knowledge or does it merely transmit knowledge? And are our testimonial belief forming practices flawed in any way (for example, by unfairly disadvantaging certain groups, or by being unreliable)? The formulation of these questions often assumes that the recovery of asserted contents is relatively unproblematic. For example, questions concerning testimonial knowledge or justification might be formulated as follows: a speaker S has said that *p*, and an audience recognises that the speaker has said that *p*, what further conditions (if any) must be in place for the audience to be justified in believing *p*, or for them to gain knowledge that *p*? Very little attention is paid to the process by which the speaker manages to say that *p*, or by which the audience recognises that *p* has been asserted. These process are generally taken for granted as reliable. A brief look at the literature in the philosophy of language illustrates that this assumption is naïve. This research reveals that audiences often rely extensively on contextual knowledge to recover the contents of an assertion, and that this processes of interpretation leaves a lot of scope for error. That is, this research reveals that there are many ways for communication to go wrong, leading to miscommunication. The aim of this dissertation is to investigate the epistemology of testimony through the lens of the philosophy of language, and to explore some of the problems which arise at the intersection of these two areas.

This introductory chapter will proceed in four sections. The first section will explore several ways in which testimonial belief formation can go wrong. These are through intentional deception, speaker unreliability, miscommunication, or environmental factors. In the second section I will give a brief overview of some of the most influential theories of testimony - theories which characterise the recent debate, in order to indicate that the issue of miscommunication (and the processes from which it arises) has received little attention. In the third section I will discuss a type of argument I call 'mismatch arguments' which appear in the philosophy of language and seem to suggest that minor miscommunications (or, the failure of speaker and hearer to coordinate on precise propositional contents) might actually be a a very common occurrence. In the final section I will

outline some recent research from social psychology which seemingly reveals both new sources of miscommunication, and that speakers and hearers may often over estimate their ability to coordinate with one another reliably.

Together the aim of these four sections is to show that there is an important gap in the current research on the epistemology of testimony, and that this gap might be fruitfully investigated by looking at the implications of recent work in the philosophy of language for the epistemology of testimony. There is a gap because there are three ways in which testimony is regularly rendered defective, yet current theories are generally formulated with a focus on only two of these defects. Moreover, this gap seems to arise out of a general lack of focus on the processes underlying our communicative exchanges. This raises the suspicion that if there is anything interesting to say about this third type of defect, and the processes it arises out of, it is likely to have been missed. It is important to explore this type of defect because it seems to be a rather mundane and often unnoticed aspect of our communicative practices (this is illustrated by the discussion of mismatch arguments, and recent research in social psychology). Yet, as we will see in the following chapters, it can render testimonial belief defective in a number of ways. Such defects can be fruitfully investigated by looking at the epistemology of testimony through the philosophy of language, as the philosophy of language gives us a picture of some of the ways in which communication actually does go wrong. I close this introductory chapter with an overview of the chapters of the thesis, and a little information on how they are connected.

1. The Ways Testimony Can Go Wrong.

The issue which receives perhaps the most attention in the epistemology of testimony concerns the conditions under which we gain testimonial knowledge, or are justified in forming testimonial beliefs. The dispute primarily concerns how demanding the conditions for testimonial knowledge or justification actually are. That is, the debate concerns whether or not (and to what extent) audiences must have reasons to believe that the testimonial exchange they are involved in has not gone wrong in some way. Some theorists think that we have a default right to assume that nothing has gone wrong (at least in particular ways) (for example, Coady 1992, and Burge 1993). Others think that audiences need to posses positive reasons to think that the testimony they are receiving is reliable (for example, Fricker 1994, 1995). In order to judge the viability of such positions it is important to

be clear on the ways in which testimonial belief formation can go wrong.

1.1. Honest Unreliable Testimony.

The first way for testimony to go wrong is for the speaker to be honest but unreliable. The simplest way for this to happen is for the speaker to mistakenly take themselves to know that *p* when *p* is in fact false. In such a situation the speaker might tell an audience member that *p*, and if the audience member were to believe the speaker then that audience member would form a false belief. In general an unreliable belief former will be an unreliable informant. They will assert on the basis of false beliefs. Believing such an informant will often lead to the formation of false beliefs. It has been argued that honest false testimony can produce testimonial knowledge on occasion. For example, Sanford Goldberg (2001) presents the following case:

> 'Martha is a very reliable friend of George. George hears Martha talking about a party she attended last night. Speaking of what she saw there, Martha reports that Jones was wearing a pink T-shirt at the party. (Context makes it perfectly clear that Martha makes this judgement on the basis of an observation she made while at the party.) However, George happens to know that Jones was not at the party in question (George was privy to Jones' last-minute decision not to attend the party; Jones subsequently spent the entire evening in question with George). George also knows that there are several people in town who are often mistaken for Jones (he is a rather regular-looking fellow). Still, none of George's doubts on the score of who Martha saw warrants any doubts in his mind on the score of Martha's sincerity, nor does George doubt Martha's observation regarding the colour of the T-shirt worn by the person in question (whoever that person was). Consequently George accepts Martha's testimony in so far as this testimony includes the informational content that *someone of other* was wearing a bright pink T-shirt at the party. That is, having heard Martha's testimony, George comes to believe that someone was wearing a bright pink T-shirt at the party.' Goldberg (2001), 516.

As Goldberg notes, Martha's testimony does present as true the proposition that someone or other was wearing a pink T-shirt. However, it also presents as true the false proposition that Jones was wearing a pink T-shirt. Most audience members not in George's position would have formed the

false belief that Jones was wearing a pink shirt. Thus, although this was an instance of testimony which did produce knowledge, it very easily could have produced a false belief. Thus it still seems defective. Moreover, Martha's assertion was incapable of producing knowledge regarding the primary content of her assertion (that Jones was wearing a pink T-shirt), because that content was false. Thus, it still seems reasonable to conclude that the vast majority of cases of false (but honest) testimony are defective and incapable of producing knowledge (certainly concerning the primary contents of the testimony).

Far more controversial is the issue of whether or not testimonial knowledge can be gained from sincere true testimony that p when the speaker does not know that p. It may seem natural to conclude that no such knowledge can be gained. Testimony seems to be a way of passing on knowledge, not generating new knowledge. And if a speaker doesn't know that p despite the fact that p is true then this surely must be because the speaker's belief is unjustified or otherwise unreliable. If so then their testimony that p will be unsafe. That is, they would assert that p even if p were false. Thus, beliefs based on such testimony will not be reliable. This seems to be a good general reason for considering such testimony to be defective. However, it is once again worth noting that are cases where such testimony can generate knowledge. Cases are given by both Sanford Goldberg (2005) and Jennifer Lackey (1999). Goldberg's case is as follows:

'Frank is a writer with a strange habit. Every morning, at precisely 7:30 a.m., he wakes up and dumps out whatever is left of the pint of milk he purchased the day before, but places the empty carton back in the fridge until noon. Then, throughout the interval from 7:30 to noon, he always remains in the kitchen, as that is where he writes every morning like clockwork. Finally, at exactly noon, he takes the now-empty milk carton out of the fridge and throws it away – an act which to him symbolizes the end of his day's writing. Now Mary is unaware of Frank's milkdumping practice. One morning, having spent the prior evening at Frank's house with Frank and her son Sonny, she awakens at 7:40 and goes to the kitchen with Sonny. Upon entering (Frank is already there) she immediately goes to the fridge for a glass of OJ, and as she reaches for the OJ she casually observes a small carton of milk. She goes on to tell Sonny (who always has cereal with milk for breakfast) that there is milk in the fridge. As luck would have it, there is indeed milk in the carton on this day (Frank failed to remember that he had bought milk yesterday). When Frank observes Mary's testimony, he realizes that he forgot to dump the milk; when Sonny observes her testimony, he forms the belief that there is milk in the fridge.' Goldberg (2005), 302.

Goldberg argues that Sally does not know that there is milk in the fridge. This is because usually there would not have been milk in the fridge, yet her evidence would have been unaltered, thus she still would have believed that there was milk in the fridge. As a result, Sally is lucky that her belief regarding the milk is true. That means that she asserted on the basis of a belief which could easily have been false. However, the belief that Sonny forms on the basis of Sally's testimony does seem to qualify as knowledge. This is because in the situations where Sally's testimony was false Frank would have corrected her, meaning that Sonny would not have formed a false belief. In this context Sonny's belief that there is milk in the fridge seems to be reliable, and thus a plausible candidate for knowledge. Thus, it seems that the speaker's lack of knowledge does not in principle block the audience from acquiring testimonial knowledge. However, this does not seem to raise doubts about the claim that such testimony is usually defective, as most cases are not like this.

1.2. Dishonest Testimony.

The first way for testimony to go wrong was for an honest speaker to make an assertion on the basis of a false or otherwise epistemically defective belief. The second way for testimony to be defective is for the speaker to be dishonest. I use the term 'dishonest' rather loosely to denote cases where a speaker says that p despite not believing that p. This includes cases from clear outright lies to cases in which the speaker fails to believe what they assert, despite having no ill intent toward the audience. It is not clear that all cases which match this definition are cases of genuine dishonesty in the strictest sense[1]. It should be clear how dishonest testimony is defective. Usually if a speaker asserts that p despite their belief that p is false (or their lack of belief that p is true) their assertion will either be made because p is false (in which case it will be reliably false), or (perhaps more commonly) it will be made because the speaker wants the hearer to believe that p regardless of whether or not p is true, meaning that the belief the audience forms will be either false or merely luckily true. This gives us good reason to avoid dishonest informants in general. Once again however it is worth noting that there do seem to be cases in which audiences can gain knowledge

---

[1]

Moreover, there are cases of dishonesty not covered by the characterisation I use here. For example Jennifer Saul (2012) asks us to consider a case in which one protagonist, Frieda, has a deadly peanut allergy, and the other protagonist, George, wants to kill Frieda. In order to kill Frieda George cooks her a meal using copious amounts of peanut oil. Frieda asks George whether or not the dish contains peanuts and George truthfully states 'there are no peanuts'. Although his utterance is true the dish does contain peanut oil and it will still kill Frieda. This seems to be a case of dishonesty which is not captured by the characterisation given above.

from dishonest testimony (at least as characterised here). Jennifer Lackey provides a compelling example of testimonial knowledge from testimony the speaker believes to be false (this is, of course, also a case of testimonial knowledge from a speaker who does not know the truth of what she asserts):

'Suppose that a Catholic elementary school requires that all teachers include sections on evolutionary theory in their science classes and that the teachers conceal their own personal beliefs regarding this subject-matter. Mrs Smith, a teacher at the school in question, goes to the library, researches this literature from reliable sources, and on this basis develops a set of reliable lecture notes from which she will teach the material to her students. Despite this, however, Mrs Smith is herself a devout creationist and hence does not believe that evolutionary theory is true, but she none the less follows the requirement to teach the theory to her students. Now assuming that evolutionary theory is true, in this case it seems reasonable to assume that Mrs Smith's students can come to have knowledge via her testimony, despite the fact that she fails condition (ii) and hence does not have the knowledge in question herself. That is, it seems that she can give to her students what she does not have herself. For in spite of Mrs Smith's failure to believe and therewith to know the propositions she is reporting to her students about evolution, she is a reliable testifier for this information, and on the basis of her testimony it seems that the students in question can come to have knowledge of evolutionary theory.' Lackey (1999), 477.

Once again then it seems that there are cases in which dishonest testimony can yield testimonial knowledge. However, once again such cases seem to be the exception rather than the rule. In the majority of cases dishonest testimony will fail to yield knowledge, thus such testimony is, in general, defective.

1.3. Miscommunication.

As we will soon see the first two defects occupy a central place in contemporary discussions of testimonial knowledge. The central debate is often framed in terms of the conditions under which we are entitled to treat the speaker as an honest and reliable belief former. However, these two defects do not exhaust the ways in which testimony can go wrong. One type of defect which has been given far less attention is that of miscommunication. Yet miscommunication can be just as

damaging. Miscommunication is understood here to occur when the proposition the speaker intends to communicate and the proposition the hearer interprets the speaker as attempting to communicate differ. Of course, it is unclear just how much these contents must differ in order for the term 'miscommunication' to really apply. If the audience recovers a proposition extremely similar to the proposition intended by the speaker, a proposition such that a world at which it differed in truth value to the speaker's intended proposition would have to be very different to our own world, then we may wish to avoid describing it as an instance of miscommunication. Communication is something which can be more or less successful, so it seems that we should treat miscommunication as a gradable phenomenon. This dissertation will be explore cases from relatively major instances of miscommunication, through to cases where no miscommunication occurs at all, but where the mere possibility of miscommunication gives rise to epistemic problems.

Miscommunication usually has as its source a flaw in the processes by which we understand one another's testimony. The lack of emphasis on miscommunication seems to be partly a result of a more general lack of emphasis on the processes underlying communicative exchanges. In this thesis I aim to emphasise the importance of the processes underlying communicative exchanges for the epistemology of testimony.

Miscommunication can take several forms. For example, you might mishear someone's testimony, and thereby form a belief in a proposition the speaker never intended you to come to believe. Imagine Samantha and Alex are talking about how Samantha is going to travel to London. Samantha says 'I'm getting a cab', but Alex mishears her as saying 'I'm getting a car'. On the basis of this Alex forms the belief that Samantha is buying a car and driving to London. This belief would clearly not constitute knowledge. It is likely false, and if it were true (for example, if Samantha were lying to Alex and was in fact intending to buy a car and drive to London) then Alex's belief would be luckily true. Miscommunication can also happen when the audience doesn't mishear the speaker, but rather misinterprets them. For example, one might mistakenly interpret 'I'm going to the bank' as 'I'm going to the (financial) bank' when the speaker intended 'I'm going to the (river) bank'. Or one might misinterpret an utterance of 'every beer is in the bucket' to mean 'every beer (for the party) is in the bucket', when the speaker intended 'every beer (belonging to the audience member) is in the bucket'. Once again, in such cases the audience's belief will usually be false or luckily true[2]. In this thesis I will primarily be concerned with cases of misinterpretation, since I am

---

[2] It is not clear that misinterpretations will always produce luckily true beliefs. Case three in chapter two can be modified in such a way that the audience recovers a proposition other than what is said, yet could still plausibly be taken to achieve testimonial knowledge. Additionally, several considerations are given early in chapter three to

investigating the epistemology of testimony through the lens of the philosophy of language, and the focus in the philosophy of language is generally on interpretation. However, I do consider mishearings in chapters five and six.

Perhaps the reason that miscommunication has received so little attention is the fact that the beliefs formed are not usually beliefs in the propositions that the speaker has attested to. For example in the cases above the speaker testified that they were going to take a cab to London, that they were going to the financial bank, and that all of the audience's beer is in a bucket. Yet none of the audience's beliefs were in these propositions. Thus they might be thought not be to be properly testimonial beliefs.

While there may be some truth in this there is still good reason to consider miscommunication when studying the epistemology of testimony. This is because the beliefs we form via miscommunication are still beliefs we form whilst attempting to form testimonial beliefs. Moreover, no doubt a few of the beliefs we hold, and consider to be testimonial, are in fact beliefs based on some form of misinterpretation. Additionally, the notion that we may have misheard someone's testimony will usually act as a defeater to whatever belief we formed on the basis of that testimony. Leaving the possibility of miscommunication out of our epistemic theorising about testimony would be like leaving the phenomenon of hallucination out of our epistemology of perception. If I were to hallucinate a green apple on the table in front of me then there is a sense in which my belief that there is a green apple in front of me would not really be perceptual. After all, it is not caused by a perception of an actual green apple on the table because there is no such object. Nonetheless, it would be a mistake to leave the possibility of hallucination out of our epistemology of perception. It is likewise a mistake to leave the possibility of miscommunication out of our epistemology of testimony.

1.4. Epistemically Unsafe Environments.

I will soon provide a brief overview of the debate over testimonial knowledge/justification in order to indicate the extent to which discussion has focused on the first two defects rather than the third. However, an overview of the potential defects of testimony would be incomplete without mentioning one final form of knowledge blocking defect discussed by Jennifer Lackey (2003).

---

indicate that there may be cases of testimonial knowledge arising from testimony where agents fail to share contents.

Lackey argues that testimonial knowledge can also be blocked by environmental factors. That is, there can be cases in which a speaker asserts *p* on the basis of their knowledge that *p* and with the intent to transmit their knowledge that *p* to the audience, the audience has no reason to doubt the testimony, and forms a belief in *p* on the basis of reliable comprehension of the utterance, and yet still fails to gain testimonial knowledge that *p*. Lackey provides the following case:

> Marvin is an epistemically impeccable recipient of testimony. Not only does he have the capacity for being sensitive to defeaters, he is also appropriately and reliably sensitive to them when accepting the reports of others. One evening while on a road trip, Marvin stops in a small town to find a hotel for the night and he encounters a large group of people gathered for an annual neighbourhood parade. Out of this large crowd, Marvin quite fortuitously focuses on Alfred, the only member of this epistemic community who reliably shares information with ''outsiders,'' and asks him where they are. Alfred tells Marvin that they are in Smithville. Now, it is true that they are in Smithville, Alfred is a reliable testifier with respect to this information, and Marvin, not knowing that the testimonial habits of the members of this community are extremely unreliable when ''outsiders'' are involved, readily accepts the proffered testimony. Moreover, because he has no reason for thinking that Smithville differs from his own community in any epistemically relevant way, Marvin does not believe, or have evidence available to him such that he should believe, that Alfred's report is either false or unreliable. Lackey (2003), 714-715.

As Lackey points out this seems to be a case in which Marvin fails to gain testimonial knowledge. This is because although he received testimony from a reliable informant he very easily could have received testimony from an unreliable informant, and in such a case he would have formed a false belief. Marvin was lucky to have picked the one reliable testifier in Smithville, thus he is lucky that his belief is true. Such environmental problems are also not widely discussed. However, it is unclear that this is really a gap in the current literature, as it seems that communities like Smithville are (hopefully) rare. Thus it seems likely that very few of our testimonial beliefs are in fact defective in this way.

It is worth noting a similar type of defect which relates to the problem of miscommunication and may occur with some degree of regularity (as will be argued in chapter three). Although it seems that we are rarely lucky to have selected a reliable informant, it does seem that we are sometimes lucky in our interpretation of an informant's testimony. In some cases (especially cases of context

sensitivity) there will be multiple similar ways of interpreting a speaker's testimony. In such cases we will usually settle on one relatively determinate interpretation. Yet, when multiple interpretations are left open the processes which lead us to reach a particular interpretation will not reliably select the interpretation intended by the speaker (In a similar way to our inability to reliably select an honest informant in Smithville). Thus, in some cases in which we do interpret the speaker correctly we will only do so luckily. In such cases we could easily have misinterpreted the speaker and formed a belief on the basis of that misinterpretation. As already discussed such beliefs will often be unreliable. Thus, in a similar way to Marvin, we would be lucky to receive reliable testimony. I discuss cases like this at length in chapter three, and explain why knowledge is often blocked in such cases.

2. The Formulation of Recent Views of Testimonial Knowledge and Justification.

I have outlined four ways in which testimony can be rendered defective, and I have also claimed that the current debate in the epistemology of testimony places a great deal of emphasis on the first two types of defect at the expense of the third. I do not mean by this that the majority of epistemologists of testimony focus on analysing or studying these first two types of defect, although the notions of trustworthiness and speaker-reliability are certainly studied in their own right (for recent work on the nature and role of trust see Faulkner (2011), and Hawley (2014)). Rather, what I mean is that the most influential views in the epistemology of testimony have generally been formulated and argued for with a sensitivity to the first two types of defect rather than the third. In order to motivate this claim I will provide a brief overview of some of these views. This is not intended to be a comprehensive overview, either with respect to the views covered, or the epistemology of testimony in general[3]. Rather, the aim is to give a taste for how some of the most influential views (the views which have really framed subsequent discussion) have generally been formulated[4].

---

[3] The epistemology of testimony extends far beyond considerations of testimonial justification and knowledge, for example, in this overview I leave out the topic of epistemic injustice which is an incredibly important recent topic in the epistemology of testimony (I do discuss epistemic injustice in chapter five of this thesis however).

[4] It is important to bear in mind that not all the theorists discussed here do ignore the possibility of miscommunication in their theorising. For example, Elizabeth Fricker, the primary defender of local-reductionism about testimony, has also written on linguistic understanding and the relation audiences must bear to 'what is said' in order to gain testimonial knowledge. However, the focus here is on the particularly influential statements of the primary theories of testimonial knowledge/justification - the statements which have framed subsequent discussion. These statements have tended to be formulated with attention to the first two types of defect and not the third.

The primary debate with respect to both testimonial knowledge and testimonial justification is between the reductionists and the anti-reductionists. Reductionists maintain that in order to gain testimonial knowledge, or be justified in forming a testimonial belief, one must have reasons to take the testimony one receives to be reliable. This is usually fleshed out as having reasons to take the speaker to be honest and a reliable belief former, at least with respect to the topic of their testimony (although it has also been fleshed out in terms of the audience having a sensitivity to signs of dishonesty (e.g. Fricker (1994, 1995))). This view is called 'reductionism' because it sees the reasons given by testimony as non-basic. Testimony doesn't provide reasons to believe unless coupled with reasons concerning the reliability of the speaker. Thus, testimonial justification reduces to these further reasons. Of course, in order for the reduction to be complete these further reasons must be sufficient for testimonial justification. Thus, there is room for hybrid views (e.g. Lackey (2006)) which take these reasons to be necessary but not sufficient for testimonial justification, and thus don't fully reduce testimonial justification to these further reasons. Reductionism is contrasted with anti-reductionism which holds that testimony that *p* provides a reason to believe that *p* unless there are specific reasons to doubt the reliability of the testimony. Testimony doesn't have to be supplemented with reasons to believe the speaker in order to provide reasons for belief. This is usually fleshed out in terms of audiences having a default right to assume that speakers are honest reliable belief formers. It is called 'anti-reductionism' because it sees testimony as, in a sense, a basic source of justification. That is, although perception plays a causal role in bringing about testimonial belief it does not play a justificatory role. Testimony that *p*, according to the anti-reductionist, provides reasons to believe without needing to be further supplemented with reasons to take the testimony to be reliable[5].

Much of the current interest in testimony (especially the reductionism vs anti-reductionism debate) started with C.A.J. Coady's influential 'Testimony, A Philosophical Study'. Coady endorses an anti-reductionist view of testimony, but presents that view largely in contrast to Hume's reductionist view (or Coady's interpretation thereof), so it is worth beginning with a brief outline of Hume's view (as presented by Coady). Coady's Hume maintains that we are justified in believing testimony due to our observation of a constant conjunction between certain types of testimony (e.g. testimony concerning a particular area of knowledge such as astronomy, or testimony from a particular type of source e.g. an expert) and the world. That is, we constantly observe testimony to be true, and we infer the truth of particular instances of testimony from this general observation. We infer from the

---

[5] Chapter six provides further discussion of the different ways of formulating anti-reductionist and reductionist views of testimony. In chapter six I provide an argument against reductionist and anti-reductionist views which hold that we have a default right to trust our own understanding.

fact that we have been presented testimony that *p*, and the fact that testimony is generally reliable, to the truth of *p*. The upshot of this is that in order to form a belief in *p* on the basis of someone's testimony that *p* we must have reason to believe that our informant has been honest and reliable in the past.

Coady presents several criticisms of Hume's view. Firstly he argues that Hume's view is either viciously circular or clearly false. Hume, according to Coady, holds that our justification for testimonial belief rests on our experience of testimony usually corresponding to the facts. However, as Coady observes, 'experience' can be read as either individual observation or shared experience. It is argued that if shared experience is the correct reading then Hume's view is circular. This is because our reliance on shared experience is a reliance on the testimony of others. Thus it cannot be used to justify our reliance on testimony. If 'individual observation' is the correct reading then it appears that Hume is just mistaken. As Coady points out we are simply in no position to personally verify that testimony in general is reliable, since we are in no position to test the truth of most of the things we have learned via testimony. Of course, Hume maintains that it is really particular types of testimony, not testimony in general, that we can observe to be reliable. However, as Coady points out, it is also unclear what is meant by 'types of testimony'. The candidates Coady considers are testimony on particular topics and testimony from certain kinds of informants (i.e. experts). It is observed that we are in no position to verify the testimony of others on most topics (for example the geography of east Asia) and so we are in no position to verify that testimony on such topics is usually reliable. Similarly, we are in no position to verify that supposed experts on such topics are really experts, for doing so would require an observation of a correlation between what the supposed experts have to say about the topic, and the facts about the topic. It would require that we too be first hand experts on the topic. Thus Coady concludes that Hume's reductionist view is fatally flawed.

It is unclear whether Coady's criticism of Hume on this point is entirely fair. Paul Faulkner (1998) suggests that the conjunctions we observe between testimony and the truth are taken by Hume to support a generalisation about human nature - that humans have a propensity toward truthfulness, and that our observation of conjunctions between testimony and the truth is what lays behind a reliance on the uniformity of human nature. This reading fits well with the context of Hume's discussion of testimony. Hume's view is developed in his discussion of miracles. Hume observes that any miracle would, by its very nature, have to be a violation of the uniformity of nature caused by a supernatural being. However, according to Faulkner's reading of Hume, our basis for believing

testimony rests on an assumption of the uniformity of human nature. Thus, any testimony reporting a miracle would undermine its own support (since the uniformity of human nature is just part of the general uniformity of nature). This is not the place to take a stand on Hume scholarship. It suffices for our purposes to observe that on either reading of Hume there is a focus on honesty and speaker reliability. The questions concern whether or not people in general are honest and generally speak the truth. There is no concern here for the general reliability of our recovery of testimony, and thus our ability to reliably form testimonial beliefs. Indeed, this seems to be built into the very formulation of the problem. Coady's Hume is concerned with the reliability of testimony as a source rather than the reliability of our testimonial belief forming practices in general. The two can come apart. For example, if speakers usually form beliefs reliably and speak with honesty then they will usually only testify that *p* if *p* is true. Thus, testimony will be a reliable source. However, if we are incredibly bad at understanding each-other, and very often attribute false interpretations to speakers, then our testimonial belief forming practices may be unreliable despite the fact that testimony itself is a reliable source[6].

It is worth mentioning a second objection Coady raises with respect to Hume's project. Coady later expands upon this argument in his attempt to vindicate testimony as a source of knowledge. It is also perhaps the closest we get to an argument which places emphasis on the processes underlying communication, since this argument aims to establish that Hume's view is in tension with the very possibility of a shared language by which we are able to communicate. According to Coady, Hume's project presupposes the possibility that testimony could be completely unreliable - that there might be no correlation whatsoever between the testimony and the truth. The idea here is that since Coady's Hume thinks that the only reason we have for accepting testimony is the observation of true past testimony it must have been possible that we observe the opposite - a complete lack of connection between testimony and the truth. Coady argues that this is not possible, for the practices of testifying and acting on testimony simply wouldn't make sense in a community where this is the case. It would be impossible to discover such a community because we would never have reason to attribute to them the act of testifying, nor would it make sense to attribute meanings to their reports which consistently render those reports false. It is important to note that this point does not rest on

---

[6] This of course partly turns on when one is counted as having testified that *p*. It might be thought that if one is interpreted as having testified that *p* (and perhaps some other conditions are met) then one has testified that *p*. Views of 'what is said' along these lines exist in the philosophy of language under the heading of 'content relativism'. On such views 'what is said' is somehow audience dependent. Such views can (but needn't) be spelled out in such a way that a speaker says that *p* to an audience if the audience interprets them as saying *p* (other conditions can be added to constrain which interpretations determine what is said). However, it would be a big departure from orthodoxy to interpret Coady's Hume, and those who have bee influenced by Coady, to attribute to them the view that the facts regarding whether or not a speaker has testified that *p* very relative to an audience member.

the more controversial claim that such a community would be impossible (although Coady does later provide some arguments in support of such a claim). Rather, the idea is that even if there were such a community we would never be in a position to discover that their communicative practices were disconnected to the truth in this way. The point is purely epistemic.

On Coady's positive view testimony is treated as a direct and fundamental source of knowledge, and the reliability of testimony as a source is vindicated via an argument which builds upon his second objection to Hume. On the topic of directness Coady considers the claim that our testimonial knowledge must be based on (or mediated by) assessments of the reliability of the speaker. This seems intuitive, for we should surely avoid consuming the testimony of unreliable informants, and surely the only way to do this is by assessing the reliability of other interlocutors. Coady rejects this hypothesis for two reasons. Firstly it appears to get the phenomenology of our actual testimonial belief forming practices wrong. The fact is that we usually do seem to just accept testimony without much critical consideration. For example, Coady describes the experience of ringing the telephone company in order to inquire about a bill. Generally you will just accept the word of the anonymous voice on the other end of the line without thinking twice. No consideration of the reliability of this individual takes place[7]. Secondly, we are often not in a position to assess the reliability of our informants. For example, it might not be clear what considerations we could bring to mind in order to gain knowledge about the reliability of our telephone company informant (although see footnote five below). If our justification for forming testimonial beliefs is not mediated by such considerations though it must be direct. On the topic of fundamentality one might worry that since perception seems essential for the acquisition of testimonial knowledge it is more fundamental. Moreover, it seems that perceptual (or otherwise non-testimonial) knowledge (or belief) must be present at the start of a chain of testimony. Coady rejects the hypothesis that perception is more fundamental than testimony. He draws an analogy between testimony and inference: we would not consider inference to be any less fundamental than perception, yet it seems that we need perception to provide the premises for many of our inferences. Moreover, Coady also points out that our perception seems to be affected by our memories and testimonial beliefs. That is, our perception is theory laden (or cognitively penetrated). Thus, since testimonial beliefs contribute to and shape the content of our perceptual experiences, these experiences should not be thought of as more fundamental. Putting Coady's comments on fundamentality and directness

---

[7]   Although it does seem that consideration of the likely reliability of the source is in the background here, for you rang the phone company rather than the plumber to inquire about your phone bill knowing that they would likely be far more reliable informants with respect to this question. Moreover, if asked to justify one's trust in the anonymous voice on the end of the line most people could easily list a few reasons. For example, why would the individual lie? It is illegal for them to lie about your bill, and many of the calls are recorded, etc.

together it seems that testimony provides a direct and fundamental justification, in no need of justificatory mediation from inference or perception.

These comments on the structure of testimonial justification don't vindicate testimony as a source of knowledge however, and it is an important part of Coady's project to do so. With this in mind Coady expands on his criticism of Hume in an attempt to establish the reliability of testimony (it is not entirely clear how Coady's arguments for the reliability of testimony connect up to the actual justification and justificatory structure of testimonial belief). As noted earlier, Coady complains that Hume's view seemingly requires that it be possible for us to discover that testimony is completely unconnected to the truth. Coady rejected this view on the basis of the fact that we could not come to discover that a community's practices of testifying were completely unconnected to the truth. Coady expands on this line, drawing on Davidson (1973), and claims that in order to interpret another agent we must assume a broad background agreement, essentially as a starting point from which to make sense of the rest of their behaviour. This fact alone does not tell us much since agreement does not entail truth. However, drawing on Davidson, Coady evokes the notion of an (almost) omniscient interpreter. Such an interpreter would also have to assume a broad agreement in background belief as a stating point to interpretation. However, if all the omniscient interpreter's background beliefs are true then a broad agreement in background beliefs amounts to broadly true background beliefs for the individual being interpreted. The possibility of such an interpreter thereby appears to ensure at least some basic level of reliability. Of course this is reliability in background beliefs, not necessarily in the beliefs explicitly assented to. And, as Coady notes, the reliability garnered here is minimal, there is still room for a lot of error and unreliability. Thus, it seems this story must be supplemented. Coady takes this basic level of reliability to be buttressed by the over all coherence and cohesion of the majority of the information we receive via testimony. He expounds this as follows:

> 'It is not that we somehow 'prove' from purely individual resources that testimony is generally reliable but that, beginning with an inevitable commitment to some degree of reliability, we find this commitment strongly enforced and supported by the facts of cohesion and coherence' Coady, (1992), 173.

It is not clear to me that the apparent coherence and cohesion with the baseline of truth provided by the Davidsonian considerations is enough to establish the general reliability of testimony. It establishes that the beliefs we form by testimony are usually close enough to the truth for us to get

by, and that our testimonial beliefs do not clash in obvious ways with our more foundational beliefs. Perhaps the truth of the majority of our testimonial beliefs is the best explanation for this fact. However, I believe that it is at least theoretically possible that we often form beliefs which although false are close enough to the truth for us to get by. I believe that this situation might arise with at least some degree of regularity from certain types of miscommunication, which I explore in chapter three.   That said, it is also unclear whether the problems arising from such miscommunication are enough to raise worries about whether or not coherence can do the work Coady wants it to do.

It is worth also briefly mentioning Tyler Burge's classic (1993) statement of anti-reductionism, which bears some important similarities to Coady's. Burge, like Coady, takes testimony to provide direct justification. He draws a distinction between two different roles that perception can play in producing beliefs. Perception can play either a justificatory role or a merely causal role. Burge maintains that perception plays a merely causal role in bringing about testimonial belief. That is, perception allows us to form testimonial beliefs by enabling the presentation of propositional contents to audiences. Our entitlement to form testimonial beliefs can, according to Burge, be made sense of without any mention of perception. Burge believes that we have an a priori entitlement to accept the outputs of sources of truth, and that the very intelligibility of a message provides us with an a priori entitlement to assume that its source is a source of truth. This justifies what Burge calls the 'acceptance principle':

> '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.

Note that no inference is required here, and the role of perception is the mere presentation of propositional contents as true[8]. The justification for the acceptance principle itself is, according to Burge, purely a priori. So, like Coady, Burge takes testimony to be a direct and non-inferential source of knowledge. Burge places a lot of emphasis on facts about the systems which produce testimony - whether or not they are sources of truth. This is usually made sense of in terms of agents as sources of testimony. That is, the claim that we are entitled to assume that a system is a rational source is taken to mean that we are entitled to assume that the agents who testify are rational.

---

[8]   It has been argued by Anne Bezuidenhout (1998) that the processes by which we actually recover asserted contents are inferential, meaning that inference plays an ineliminable role in the acquisition of testimonial knowledge.

Before moving on to consider objections to the anti-reductionist view it is worth considering a final form of anti-reductionism - the Assurance view of testimony. The Assurance view is a development of anti-reductionism which places an even greater emphasis on the testifier as an honest and reliable source. It is discussed in greater detail in chapter four, so a few brief remarks will suffice at this point. Assurance theorists, like most anti-reductionists, object to the notion that testimonial justification is inferential - that we infer from generalisations about testimony, or the particular situation we are in, to conclusions about the likely truth of what is asserted. To my mind the strongest consideration offered by the assurance theorists in favour of this conclusion is offered by Angus Ross (1986). Ross points out that to form testimonial beliefs by inferring the likely truth of the testimony from generalisations about testimony one has received in the past one must take a detached view of the actions of the testifier, fitting them in with a broader pattern of generalisations about testimony in general (or particular kinds of testimony). However, to do so would not be to treat the testimony in the spirit in which it is offered. Indeed, as Ross points out, it is hard to see how an asserter could view their own testimony in such a detached way.

Rather, when we assert we seem to give our word to the audience, or offer an assurance of the truth of the proposition attested to. In doing so, according to the assurance theorist, we take responsibility for the audience's testimonial belief. It is this, for the the assurance theorist, which is the source of our right to form testimonial beliefs. Instead of inferring from generalisations about testimony as a source, or even the likely truth of the proposition attested to, we simply accept the speaker's offer of assurance, and in doing so pass on the responsibility for the truth of that belief to the speaker. There is a strong emphasis on the relations of trust between the audience and speaker here - trust in the speaker to take on responsibility for the content offered as true. However, there is no consideration of the processes by which contents are actually offered as true, or by which audiences are able to judge or understand that a particular content has been offered as true. In chapter four it is argued that even the assurance theorist, with their strong emphasis on the interpersonal aspects of testimonial exchanges, can benefit from a consideration of the processes underlying the recovery of propositional contents. This is because the realistic possibility of miscommunication allows speakers to back out of the commitments they undertake when asserting by claiming that they have been misunderstood.

This concludes the brief overview of the anti-reductionist approaches. I now turn to Elizabeth Fricker's objections to the anti-reductionist program, and her local-reductionist alternative. Fricker

(1995) frames the question of testimonial knowledge as follows: there are certain conditions which must obtain for a source of belief (such as perception or testimony) to be functioning properly (call these '*V*-conditions', or 'veracity conditions'). The *V*-conditions for testimony are, according to Fricker, that the speaker is honest, and that what she asserts is true (in later papers Fricker develops her view in terms of the speaker knowing that what she asserts is true, rather than it simply being true). These conditions correspond to the first two testimonial defects discussed above (unreliable belief formation and dishonesty). So long as these conditions obtain, Fricker asks, what other conditions (if any) must be met in order for the audience to gain testimonial knowledge? We could either side with the reductionist and take it that the audience must have some independent reasons to hold that the *V*-conditions obtain, or we could side with the anti-reductionist and conclude that no such additional reasons are required - as long as these conditions obtain the audience has a right to form a testimonial belief. Fricker agrees with Coady that Humean global reductionism is a non-starter. However, she also objects to the notion that we have a presumptive right to believe *p* simply on the basis of the fact that *p* has been asserted. To form beliefs in such a way would, according to Fricker, be to exemplify a form of gullibility. It would amount to uncritical acceptance of whatever one was told unless one had good reasons for rejection.

Fricker's alternative is what she calls 'local-reductionism'. Local-reductionism does not require that one make inferences from the general reliability of testimony. Indeed, local-reductionism does not require that any specific inferences about the reliability of the testimony being offered be made either. What local-reductionism does require is that the audience be on the look out for signs of incompetence with respect to the subject-matter, or untrustworthiness. Fricker argues that we do have a right to assume that others are competent and honest with respect to certain subject matters (for example, their name), but in order to avoid gullibility we must be sensitive to signs of dishonesty or incompetence. It is not clear what this monitoring requirement amounts to in practice. One might not be blamed for thinking that the requirement is that audiences consciously be on the lookout for signs of dishonesty. This would be at odds with Coady's claim about the phenomenology of testimonial belief formation. However, as Goldberg and Henderson (2006) point out our sensitivity to signs of dishonesty might be largely subpersonal (Goldberg and Henderson take such subpersonal monitoring to be consistent with anti-reductiomism. Whether or not such a view would be properly described as reductionist or anti-reductionist is unclear, this issue is discussed futher in Fricker's (2006b) response to Goldberg and Henderson)[9]. More recently Fricker

---

[9] A similar view to Fricker's is endorsed by Sperber et al (2010), who maintains that a faculty for 'epistemic vigilance' toward dishonesty plays an important role in explaining the evolutionary stability of our communicative practices. Kourken Michaelian (2010) responds, arguing that work on lie detection in empirical psychology demonstrates that

(2006a) has developed her view in a similar way to the assurance theorist, with a focus on the speaker taking responsibility for the propositions they assert. However, unlike the assurance theorist Fricker still takes testimony to be a form of evidence. The fact that speakers take responsibility for what they assert ensures, according to Fricker, that assertions (or a limited subset of assertions called 'tellings') are governed by the knowledge norm of assertion. Thus, when audiences are within their rights to assume that the speaker is trustworthy they are in their rights to assume that the speaker knows the truth of what they assert (not merely that it is true). This development of Fricker's view will be discussed in chapter four.

Finally, a third approach to testimonial justification is offered by Jennifer Lackey (2006, 2008) who brings together elements of reductionism and anti-reductionism to advocate a hybrid view. Lackey endorses the reductionist's necessary condition on testimonial justification - that in order for the audience's testimonial belief to be justified they must possess reasons to consider the testimony reliable. However, as Lackey notes, the necessary condition alone is not sufficient for a reduction. Testimonial justification is only reduced to these positive reasons if the positive reasons are sufficient for testimonial justification. Lackey denies that they are sufficient. She presents the following case which purports to illustrate that one can have positive reasons to accept testimony without thereby being justified in doing so:

> 'Nested speaker: Fred has known Helen for five years and, during this time, he has acquired excellent epistemic reasons for believing her to be a highly reliable source of information on a wide range of topics. For instance, each time she has made a personal or professional recommendation to Fred, her assessment has proven to be accurate; each time she has reported an incident to Fred, her version of the story has been independently confirmed; each time she has recounted historical information, all of the major historical texts and figures have fully supported her account, and so on. Yesterday, Helen told Fred that Pauline, a close friend of hers, is a highly trustworthy person, especially when it comes to information regarding wild birds. Because of this, Fred unhesitatingly believed Pauline

---

our ability to monitor others for signs of dishonesty is highly unreliable.

earlier today when she told him that albatrosses, not the widely believed condors, have the largest wingspan among wild birds. It turns out that while Helen is an epistemically excellent source of information, she was incorrect on this particular occasion: Pauline is, in fact, a highly incompetent and insincere speaker, especially on the topic of wild birds. Moreover, though Pauline is correct in her report about albatrosses, she came to hold this belief merely on the basis of wishful thinking (in order to make her reading of *The Rime of the Ancient Mariner* more compelling).' Lackey (2006), 163.

Lackey argues that despite the fact that Fred has very good reasons for accepting Pauline's testimony his belief was not justified, because Pauline is not in any way a reliable source. If Lackey is correct then positive reasons for accepting testimony are not sufficient for testimonial justification, meaning that reductionism is false[10]. Lackey postulates that in order for a testimonial belief to be justified the speaker's testimony must also be reliable. More precisely, she specifies her view as follows:

'Dualism: For every speaker A and hearer B, B justifiedly believes that *p* on the basis of A's testimony that *p* only if: (1) B believes that *p* on the basis of the content of A's testimony that *p*, (2) A's testimony that *p* is reliable or otherwise truth conducive, and (3) B has appropriate positive reasons for accepting A's testimony that *p*.' Lackey (2006), 170 .

The first condition here is intended to rule out cases in which one simply infers a truth from some fact one observes about the testimony. For example, if someone asserts something in a soprano voice then you might infer that they have a soprano voice. You have not thereby gained testimonial knowledge that they have a soprano voice. The second condition is spelled out in terms of either the

---

[10] As a side point, it is not clear that Lackey is correct about Fred's belief being unjustified. This seems to be a gettier case - Fred's belief is justified but luckily true. Lackey considers and rejects this hypothesis, maintaining that the fact that Pauline is in no way reliable (rather than usually reliable, but unreliable on this occasion) bars this from being classed as a getter case. This strikes me as a weak response. Consider the following modified version of Russell's stopped clock case:

Fred is visiting his parents, who have just refurbished their house. His parents (who are usually extremely reliable and honest) have decided to play a joke on Fred - they have purchased an incredibly convincing replica clock which always shows the time 3:30, and placed it in their lounge. Furthermore, they have told Fred that they have purchased an incredibly accurate new clock and that they have placed it in the lounge. Fred doesn't happen to enter the lounge and see the replica clock until exactly 3:30. When he does enter the room he sees the replica and forms the true belief that it is 3:30.

This case seems directly analogous to Lackey's nested speaker case, since the replica clock is in no way a reliable source with respect to the time. It also seems to be a clear gettier case. Fred has excellent reasons to believe that the replica clock is in fact a real reliable clock. Thus, his belief that it is 3:30 seems to be justified. Moreover, his belief is true, yet it still fails to constitute knowledge.

speaker being honest and a reliable belief former, or the speaker being disposed to make the assertion she makes only when the content of that assertion is true. Lackey lists three types of positive reason the audience might have in order to satisfy the third condition (other than past experience with the particular speaker's true testimony on the particular topic of the testimony). The audience is likely to have reasons for considering certain types of context more reliable, they will know that people are usually reliable and honest with respect to certain topics (e.g. their own names), and they will know that certain types of informant are reliable with respect to certain topics (for example accountants are reliable with respect to financial matters). Once again the focus here on the reliability of informants (in terms of both honesty and expertise), not the communicative process itself.

I have outlined a selection of views concerning testimonial knowledge and justification, and emphasised the fact that these views are usually formulated with a sensitivity to the first two forms of testimonial defect, and not the third. In Coady we saw a deep concern for our ability to gain inferential knowledge concerning the truthfulness of much of the testimony we receive (which is, of course, an extremely important topic in its own right), but no consideration of our right to assume that we reliably recover the correct contents. With the assurance theorist (and Fricker's later work) we saw an emphasis on the commitments a speaker undertakes when making an assertion (an element of honest testimony), and the role these commitments play in entitling us to form testimonial beliefs. In Fricker we saw an emphasis on monitoring for signs of untrustworthiness and incompetence with respect to the topic of the testimony, and in Lackey we find a discussion of the types of factors we are sensitive to with respect to the honesty and truth conduciveness of the testimony we receive. However, there is no mention of monitoring for miscommunication, or sensitivity to signs thereof. Indeed, we even saw that the very formulation of these views in terms of the reliability of testimony, rather than the reliability of our testimonial belief forming practices, lent itself to such a lack of focus on the processes by which we actually recover the contents of testimony.

This doesn't necessarily amount to a deep flaw in the debate. Many of these views can no doubt be reformulated to make explicit their assumptions about the extent to which speakers and hearers reliably communicate, or the extent to which communication must be reliable in order for audiences to gain testimonial knowledge. For example, the local reductionist or hybrid theorist might require that we not only be sensitive to signs of dishonesty or incompetence, but that we must also be sensitive to signs of miscommunication. Likewise, the anti-reductionist might consider

miscommunication to be a defeater which blocks testimonial knowledge. It might even be thought that if miscommunication occurs then the audience's belief is not properly testimonial (since in cases of miscommunication the proposition recovered is not the proposition the speaker attested to). Moreover, this lack of attention is not especially worrying if there is little to be said about miscommunication, if miscommunication is rare, or if we reliably identify instances of miscommunication and avoid forming beliefs when miscommunication has occurred. In the remainder of this introductory section I will turn to the philosophy of language and social psychology to illustrate that communication is not always as smooth as we might hope. The aim primarily being to establish that miscommunication is a relatively frequent occurrence, but also to indicate the sorts of factors it arises out of, and the fact that it may often go unnoticed. In the remainder of the thesis I explore several different issues which arise from consideration of the processes underlying our communicative exchanges, particularly with respect to miscommunication.

3. Context Sensitivity and Mismatch.

So far I have described four ways in which testimony can be defective, and I have given a brief overview of some of the most influential views of testimonial knowledge and justification in order to illustrate that they are generally formulated with a sensitivity to the trustworthiness and reliability of the speaker, rather than the reliability of the processes by which we recover communicated contents. I also noted that miscommunication (which I characterised as arising whenever the proposition the speaker intends and the proposition the audience recovers differ) often seems to block testimonial knowledge. This would not be of interest if miscommunication were an uncommon phenomenon. However, work in the philosophy of language seems to suggest that miscommunication (at least fairly minor instances thereof) might be very common. This is illustrated in a series of arguments concerning our ability to successfully communicate with context

sensitive terms.

Intuitively, communicative success occurs only when the proposition the speaker intends to communicate is identical to the proposition the hearer recovers. A number of theorists have been led to reject this condition on communicative success on the basis of what I will call 'mismatch arguments'. These arguments have lead some philosophers to conclude that recovery of the speaker's intended meaning is not required for communicative success (Sperber and Wilson (1986), Bezuidenhout (1997), Heck (2002), Recanati (2004), and Carston (2002)), and others, in an attempt to maintain the hypothesis that communicative success requires recovery of the speaker's intended meaning, to reject the view that propositions are the objects of speaker meaning in the first place (Buchanan (2010))[11]. There are two versions of the mismatch argument. The first draws our attention to the seemingly miraculous nature of recovery of the speaker's intended meaning in instances of context sensitivity. The second draws our attention to cases where it is unlikely that the speaker and hearer are even able to entertain the same propositions. I will return to the arguments and views presented in this section repeatedly throughout the thesis, so to avoid unnecessary repetition I will keep this overview relatively short[12].

The first version of the mismatch argument has its roots in the debate over definite descriptions, going back to Wettstein's (1981) argument against the Russellean theory of definite descriptions. Wettstein noted that in most cases in which we attributively use a definite description, the explicitly articulated description is not sufficient to identify a unique referent. Consider the sentence 'the book is red' - there are many books in the world, indeed there are many red books, thus the sentence alone is not sufficient to identify the unique red book intended. It must seemingly be supplemented in order to uniquely designate a particular book. However, there will be many ways in which it can be supplemented in order to uniquely designate a particular book. Wettstein asks us how Russell's theory of descriptions can deliver the result that an utterance of such an 'incomplete' definite description can possibly be determinate, and he offers us an alternative theory of definite descriptions on which the content is determinate. However, it has since been observed that similar problems arise for a large number of context sensitive terms. For example, Schiffer (1992) notes

---

[11] In this dissertation I am not concerned with the conditions of communicative success. Indeed, I am not convinced that much can be gained from studying communicative success since I think it is an extremely vague target of inquiry. Success is always success at some task - at something with a particular aim. However, it is not clear that we have a single aim when we communicate, or that 'communication' picks out a single clear task. Speakers, hearers, and the interlocutors as a group may all have multiple separate aims in typical instances of communication. This is one reason for focusing on testimonial knowledge rather than communicative success - it provides a far clearer target for inquiry.

[12] For example, chapter three contains a discussion of Buchanan's recent argument that propositions are not the objects of speaker meaning, and chapters four and five discuss relevance theory. For this reason I do not go into much

that it arises for belief reports, and Buchanan (2012) notes that it arises for quantifiers and non-sentential assertions (those are the examples he discusses in detail, however he also maintains that such problems arises for most context sensitive terms).

The general problem is that in many cases in which a context sensitive term is used there will be many very similar semantic values which can be assigned to that term, each determining a different proposition. In the case of definite descriptions the sentence in need of supplementation may be 'the book is red', and two of the competing candidates might be 'the book in this room' or 'the book I wish you to read'. In the case of quantifiers the sentence could be 'every beer is in the bucket', and the competing values could be 'every beer we brought this afternoon is in the bucket' or 'every beer we purchased from the corner shop is in the bucket'. It will often be impossible to tell, without further investigation, which precise proposition the speaker intended to communicate in any one of these cases. The evidence provided by the speaker's utterance and the context will rarely be sufficient to grant the audience knowledge that the speaker intended one of these propositions rather than the other. Moreover, this will be the case for any use of a context sensitive term where there are very similar contextual values competing to be assigned - especially in cases where the values are complex and fine grained. In the cases of quantifiers and descriptions the relevant contextual values were properties. However, values such as comparison classes, standards of taste, and ordering relations on worlds will generate similar problems. Thus it seems inevitable that this phenomena arises for a very large number of utterances.

Indeed, according to some views (e.g. Sperber and Wilson (1986), Carston (2002), and Recanati (2004)), recovery of what is said always involves pragmatic processing. On such views we employ pragmatic processing to select the most contextually appropriate concept to assign to a term even when that term is not standardly thought to be context sensitive. Competing concepts are weighed against each other and the concept with the highest level of activation, or the optimally relevant concept (where relevance is a measure of relevant 'cognitive effects') is selected and assigned[13]. If such views are correct then there will be the same epistemic uncertainty in almost all cases of verbal communication, since the concepts assigned by the hearer and speaker could always be very similar but distinct. Such theorists usually adopt the view that communicative success requires only that the hearer recover a proposition similar to that intended by the speaker.

The second type of mismatch argument arises only for Fregean views of de re utterances. According to such views, in order to understand certain de re utterances one must come to entertain

---

[13]   Chapters four and five contain more complete outlines of these views.

a proposition containing a mode of presentation. However, the way the speaker and hearer conceive of the object of the utterance is likely to differ, since they occupy different perspectives, and have different knowledge of the object. Thus, the proposition the audience recovers will almost always contain a mode of presentation other than the one the speaker intended. This argument has been pressed by Bezuidenhout (1997) and Heck (2002). Heck, concerned with demonstratives, puts the first point as follows:

> '..utterances of demonstrative expressions can differ, in ways relevant to understanding and communication, even when they do not differ in their referential properties: the successful communication of information from one speaker to another depends not only upon speakers' identifying the right objects as the demonstrata, but also upon their thinking of these objects in the right sorts of ways, although it does not depend upon their thinking of the objects in any particular ways.' Heck (2002), 25.

Heck asks us to imagine the following case: we are in a long room with two windows at either end of the same wall. Out of one window we can see the bow of a ship, out of the other we can see the stern. As it happens they are the stern and bow of the same gigantic ship, but we do not know this. I gesture toward the window through which we can see the stern of the ship and say "That ship is an aircraft carrier". In order for you to understand my utterance it is not sufficient that you come to entertain the singular proposition attributing the property of being an aircraft carrier to the indicated ship. That is because if you had mistakenly taken me to be gesturing toward the window through which we can see the bow of the ship, you would have come to believe the same singular proposition. Yet in this case you would have misunderstood me, since neither of us know that the ships we can see out of either window are identical. Thus, you can only understand my utterance by entertaining a proposition containing some mode of presentation of the ship. However, it is not clear that we must grasp a proposition containing precisely the same mode of presentation as that intended by the speaker. Indeed, since we both occupy different perspectives it is unlikely we will be in a position (at the time of utterance) to entertain precisely the same mode of presentation. Sperber and Wilson (1986) make a similar point:

> '..two people may be able to think *of* the same man *that* he has gone, without being able to think exactly the same thought, because they might not individuate him in exactly the same way. Similarly, by saying 'He has gone' I may induce in you a thought which is similar to mine in that it predicates the same thing (that he is gone) of the same individual, but which differs from mine in the way you fix the reference of 'He'. It seems to us neither

paradoxical nor counterintuitive to say that there are thoughts which we cannot exactly
share, and that communication can be successful without resulting in the exact duplication
of thoughts in communicator and audience. We see communication as a matter of enlarging
mutual cognitive environments, not of duplicating thoughts.' Sperber and Wilson (1986)
192-193.

Heck points to the indexical 'I' as a clear case where the speaker and hearer cannot possibly share
the same mode of presentation:

'The problem is most dramatic in the case of 'I'. The belief that someone expresses when she
says "I am a philosopher" is the self-conscious belief that she herself is a philosopher. But
the belief I form, if I accept what she says as true, is not the self-conscious belief that she is
a philosopher: *I* cannot so much as entertain that belief. Her self-conscious belief that she is
a philosopher, though it involves a thought to which I can refer, one I might well know her
to believe, is in that sense private to her.' Heck (2002), 20-21.

If these theorists are correct then in some cases it will be impossible for speaker and hearer to share
the same proposition because they will possess different modes of presentation of the object of that
proposition. The two forms of mismatch argument taken together suggest that miscommunication
(at least, a minor form thereof where similar but non-identical contents are entertained) may be a
very common occurrence. It might be thought that the type of miscommunication discussed here is
benign enough to be easily ignored. After all, it is still taken by many philosophers of language to
be compatible with communicative success. Indeed, I believe that such miscommunication is often
benign. However, as I will argue in the main body of the thesis, I think it can also lead to several
interesting epistemic problems. Before outlining these problems though I will briefly outline a
recent series of psychological studies which also seem to reveal that miscommunication may be a
more common occurrence than we usually assume. These studies further corroborate the claim that
miscommunication is a relatively common occurrence, which often goes unnoticed, and they are
referred back to at various points in the thesis.

4. Egocentric Processing in Utterance Interpretation.

Recent experimental research has suggested that the application of theory of mind in utterance

interpretation is not always immediate and automatic, and that adults initially interpret egocentrically (appealing to their own private knowledge rather than focusing specifically on common ground information) before correcting this egocentric interpretation by application of theory of mind. However, it has also been suggested that this second stage of processing in which theory of mind is applied is not perfectly reliable, and can be blocked by a number of factors. This leads to an 'egocentric bias' whereby audiences settle on their initial egocentric interpretation, failing to apply theory of mind to check that their initial interpretation is correct. A consequence of this is that in the circumstances in which the second stage of processing is blocked or otherwise interfered with the audience will not be sensitive to whether or not they have interpreted the speaker correctly, meaning that even when they do interpret correctly they would have settled on the same interpretation even if it were not correct (that is, their correct interpretation is rendered unsafe).

Worryingly, the factors which block this second stage of processing are not the sorts of factors we would usually take to affect whether or not we have testimonial knowledge. For example, it has been shown that we commit far more egocentric errors when interpreting close friends and relatives than when interpreting strangers, and we commit far more egocentric errors when in a happy mood than when we are sad.

This two stage approach to interpretation (known as the 'Egocentric Anchoring Approach'), and the phenomenon of of egocentric bias, is motivated by a series of studies carried out in particular by Boaz Keysar (with a number of collaborators). It has been believed for a long time that young children fail to properly take the knowledge of others into account when predicting their actions. For example, if they are given a piece of information which is not available to another agent they will predict the agent's actions as if the agent possessed the child's private piece of knowledge (see Wimmer and Perner (1983), and Olson and Torrance (1987), Wellman et al (2001) provides an overview and meta-analysis). It was suggested by Keysar (1994) that this egocentric cognition actually continues into adulthood, not being straightforwardly replaced by theory of mind[14].

In order to indicate that egocentric processing continues into adulthood Keysar presented two groups of adult participants with the following case - Jane has told David that a particular restaurant is good. David then visits the restaurant. In one version of the case David likes the restaurant, and in the other he hates it. In both cases he leaves a note to Jane saying 'I went to the restaurant - it

---

[14] For evidence that young children are able to pass non-verbal versions of the false belief task see Onishi and Baillaregon (2005), He et al (2012), and Scott and Baillaregon (2013). It is not clear yet what consequences these results have for the egocentricity hypothesis.

was marvellous, just marvellous'. In one case he is sincere, in the other sarcastic. In each case participants were asked whether Jane would sense sarcasm in the note. Surprisingly, people were far more likely to say Jane would sense sarcasm when they knew that David was being sarcastic. If the participants took Jane's perspective into account however there should have been no difference in judgements about the cases. This suggests that egocentric processing is present in adults as well as children (further evidence of egocentric bias is found in Ross, Greene, & House (1977), Dunning, Meyerowitz, & Holzberg (1989), Alicke (1993), Gilovich, Medvec, & Savitsky (2000), and Epley, Savitsky, & Gilovich (2001)).

A series of experiments were then conducted to determine the effect of this egocentric processing on our language processing. On the speaker side of things Horton and Keysar (1996) found that when constructing utterances without time constraints speakers took audience information into consideration, but regularly relied on private information unavailable to the audience when under a time constraint. This suggests that in utterance planning speakers initially plan egocentrically, drawing on private information, and then adjust that plan sequentially to take into account the audience's information.

More importantly for our interests similar results were obtained when looking at the audience's processing of utterances. Two relevant studies were carried out. In the first (Keysar, Barr, Balin and Paek (1998)) participants were given two sentences describing simple events involving male or female actors. One sentence gave shared information, the other was private. The genders of the actors were sometimes the same and sometimes different. In each case the participants are asked something about the actor in the shared case. Responses took longer, and there were more errors, when the pronoun used in the question corresponded to the gender of the agent in the private case (that is, in cases where the gender of the actor in the private case was identical to that of the actor in the shared case). This was taken to show that egocentric interpretation interfered with interpretation and required a perspective adjustment. The thought being that the responses took longer because when assigning a referent to 'he' or 'she' the participants had two candidates to choose between, one from the private case and one from the shared case. If we initially only appeal to common ground information then the only available candidate would be the individual from the shared case and there would be no interference.

In a second study (Keysar, Barr, Balin, and Brauner (1996)) two participants were placed on the opposite sides of an array of objects on a series of shelves. One of these participants (a confederate)

was to direct the other participant to move objects to different locations on the shelf (call this second participant the instructee). However, some of the objects were hidden from the confederate. This fact was known by the instructee. The instructee's eye movements were tracked in order to determined which object he or she looked at first when given a particular command. It was found that when the confederate's description of an object fitted an object visible only to the instructee better than one mutually available the instructee reliably directed their gaze toward the privately available object, sometimes even causing them to reach for it, before correcting themselves and reaching for the intended object (call this experiment the 'reference task'). This too suggests that audiences initially and automatically interpret egocentrically, applying theory of mind later to correct errors in interpretation. Epley, Morewedge, and Keysar (2004) replicated these results performing the test on children and adults. They found that children and adults were just as likely to initially focus their attention on the hidden objects, but that adults were far more likely to correct themselves and select the public object meeting the description.

These findings alone are not problematic. However, as I explained at the beginning of this section, further studies have shown that the application of theory of mind in utterance interpretation is sometimes blocked by a number of factors such as mood and relationship with the speaker. For example Savitsky, Keysar, Epley, Carter, and Swanson (2010) found that egocentric processing is more prevalent when communicating with people we know well than it is when communicating with strangers (similar findings have been presented by Van Boven, Kruger, Savitsky & Gilovich (2000), and Vorauer & Cameron (2002)). They constructed a version of the reference task were some participants were paired up with friends, and others were paired up with strangers. It was found that audiences committed a far greater number of errors (for example moving the hidden object rather than the public object) when paired with a friend than when paired with a stranger. Participants were also slower to adjust their initial egocentric interpretation under these conditions. Another experiment was then carried out in which participants were paired up with strangers or friends and asked to communicate ambiguous messages. The speakers were given several ambiguous phrases with their meanings, and one meaning highlighted for each. They were asked to utter the phrase and try to convey the highlighted meaning. The speakers were then asked to judge whether the audience understood them or not. It was found that audiences who knew the speaker were only marginally more reliable interpreters than those paired with strangers. However, speakers judged themselves to be far more reliable communicators when paired with friends than when paired with strangers, significantly overestimating their effectiveness as communicators. These two studies are taken to show that we relax our application of theory of mind when

communicating with friends and relatives.

It has also been suggested that egocentric processing is more prevalent when we are happy than when we are sad. Converse, Lin, Keysar, and Epley (2008) repeated the reference task with happy, sad, and neutral participants. It was found that sad participants were marginally more reliable than neutral participants, but happy participants were significantly less reliable than neutral participants, both committing more errors and taking longer to adjust their initial egocentric attention on the hidden object. Reliability at the reference task has also been shown to be sensitive to working memory, suggesting both that we are more egocentric when under a heavy cognitive load, and that theory of mind application can be an effortful process (Lin, Keysar, and Epley (2010)).

This suggests that there is a range of circumstances in which our processing is prone to egocentricity, and the application of theory of mind to adjust our initial egocentric interpretation can be blocked. Even more worrying is that the circumstances under which this occurs are circumstances we would not normally take to be detrimental to our acquisition of testimonial knowledge. They include circumstances in which we are happy, those in which we are interacting with friends, and those in which our working memory is negatively effected[15]. The consequence of this is that we are more prone to egocentric biases under these conditions. Thus, we are less reliable at coming to know what we have been told in such circumstances (even if we are no less reliable at recovering the intended content in such cases). This is because when our initial egocentric interpretation gets things right (as it often will) it will often be the case that we would have retained that interpretation even were we to have been mistaken.

Of course, in cases like the reference task where we are interacting with the local macro level environment mistakes will quickly be corrected by the speaker. As a result audience's beliefs about the speaker's intentions will arguably still be safe and sensitive. Had the speaker intended something else then the audience would not have retained their initial egocentric interpretation because they would have been corrected (this has a similar structure to the cases presented by

---

[15] This final condition may seem less worrying than the others, because it seems fairly reasonable to claim that we are less reliable communicators when under a heavy cognitive burden. However, our working memory can be effected by a range of factors which we would not normally take to be detrimental to our ability as communicators. For example, a series of studies carried out by Jennifer Richeson has found that executive function is negatively effected by our adjusting for implicit racial biases (Richeson & Shelton (2007); Richeson, Trawalter, & Shelton (2005); Richeson & Shelton (2003); Trawalter & Richeson (2006); Trawalter & Richeson (2008); Trawalter, Richeson, & Shelton (2009)). If this is correct then agents adjusting for implicit racial bias are likely also to suffer from egocentric bias in interpretation, and thus have diminished knowledge of speaker intentions. See Tamar Szabó Gendler (2011) for a discussion of the epistemic import of these findings in which she suggests that there just are real epistemic costs to adjusting for implicit racial bias.

Goldberg (2005), discussed in section 1.1). However, this will not be the case for more fine grained contextual values such as those mentioned in the previous section (for example properties for descriptions and quantifiers, comparison classes, ordering relations etc.). In these cases as long as the egocentric interpretation is at least relatively close to the speaker's intended interpretation the initial egocentric interpretation will not usually be corrected in nearby worlds where it is false.

Keysar et al's findings have not gone unchallenged. Related studies by, for example Hanna et al (2003) have provided evidence against the claim that initial interpretation is wholly egocentric, suggesting that although egocentric perspective does play a role in initial reference assignment knowledge of the common ground does as well. Additionally Rubio-Fernandez (2008) points out that many of Keysar et al's findings are consistent with a parallel processing view where egocentric and common ground interpretations are processed in parallel and compete against one another. These responses might undermine the egocentric anchoring approach to interpretation, however they do not undermine the phenomenon of egocentric bias. That is, they do nothing to undermine the experimental results which appear to show that in certain circumstances the egocentric interpretation is rendered dominant. Thus, these responses to Keysar et al do not call into question the claim that our interpretations are, not infrequently, rendered unreliable by egocentric bias.

I take the existence of egocentric bias, together with the mismatch arguments discussed in the previous section, to suggest that miscommunication is a reasonably common phenomenon. Moreover, it appears that miscommunication involving certain types of contextual value (fine grained values such as properties or ordering relations on worlds) might often go unnoticed, as the differing interpretations they determine will often be of little practical consequence. This second claim is further supported by the studies which appear to indicate that speakers overestimate their own ability to reliably communicate ambiguous information. Thus, although miscommunication has received little attention from epistemologists of testimony, it appears to be a reasonably common phenomenon which often goes unnoticed. Thus, it is worth investigating the epistemology of testimony with an eye to miscommunication and the processes from which it arises (in particular, the processes underlying our ability to communicate with context sensitive terms). It is to this task which I now turn.

5. Overview of Chapters.

The remainder of this thesis consists of a collection of essays exploring problems which arise in the epistemology of testimony once we consider the processes underlying interpretation, and the ways in which they can go wrong. These essays, although connected by a central theme, are independent, and can be read separately.

Many of the problems which seem to arise when we consider the processes underlying our recovery of communicated contents concern the hearer's ability to reliably recover the correct meaning - their ability to gain knowledge of what is said. It is often thought that testimonial knowledge requires knowledge of what is said. If this is the case then it might be thought that an investigation into the consequences of such problems is hardly worth conducting, since the consequences are obvious - we clearly do not gain testimonial knowledge in cases in which miscommunication is an epistemic possibility. In chapter two I argue that testimonial knowledge does not require knowledge of what is said. I lay out a set of assumptions about the conditions on knowledge, and the determination of what is said, and then present a series of cases which, on the given assumptions, constitute cases of testimonial knowledge without knowledge of what is said. I consider a number of responses to these cases, and indicate at the end how analogous cases can be constructed on different assumptions about the conditions for knowledge and the determination of what is said.

In chapter three I begin the investigation into the epistemic impact of possible miscommunication. I start by suggesting that actual minor miscommunications do not always block testimonial knowledge (this is a step beyond the conclusion of chapter two which merely argues that testimonial knowledge does not require knowledge of what is said), but I spend the majority of the chapter explaining a sense in which they often do block such knowledge. Moreover, I argue that such miscommunications are reasonably widespread and may often go unnoticed. I do this by outlining a form of mismatch argument according to which the factors which give rise to mismatch are common features of many low stakes contexts. I then consider and reject a number of responses. I conclude that a mild scepticism about much low stakes context sensitive testimony may be warranted.

The epistemic significance of miscommunication (or the possibility thereof) does not end with its ability to block testimonial knowledge however. In chapters four, five and six I explore three further types of epistemic problem which are revealed by considering miscommunication (and its sources). In chapter four I argue that the possibility of miscommunication gives speakers plausible deniability about a wide range of assertions, even when miscommunication has not occurred. That is, speakers often have scope to claim that miscommunication has occurred even when it has not.

By doing so they are able to avoid undertaking robust commitments when asserting. I explore the theoretical consequences of this for theories of testimony which place a particular emphasis on the commitments speakers undertake when asserting (for example, Assurance theories, and Elizabeth Fricker's recent development of her reductionist views).

One of the most important recent developments in the epistemology of testimony has been Miranda Fricker's discussion of epistemic injustice. Fricker's approach bears significant similarities to the approach taken in this thesis. This thesis explores the consequences of defects in the processes underlying interpretation. Fricker explores the ethical and epistemic consequences of a particular type of defect in the processes underling our credibility judgements. She argues that these processes rely on heuristics and stereotypes which can encode prejudicial judgements of the speaker. This can lead to certain groups being unfairly assigned a low credibility, and gives rise to a number of harms. For example, such speakers are silenced, and disrespected as knowers. Chapter five explores a similar phenomenon which arises at the level of interpretation. Like our credibility judgements our interpretations are guided by social stereotypes, and this can lead to similar ethical and epistemic problems. That is, they can lead to certain groups being disproportionately subjected to harmful forms of misinterpretation. I explore the causes of this phenomenon, its harms, and its relation to epistemic injustice as characterised by Fricker.

Chapter six explores the implications for cases such as those discussed in chapter five for testimonial anti-reductionism. The chapter begins by outlining a recent argument presented by Susanna Siegel against perceptual dogmatism. Siegel presents cases in which an agent's irrational belief cognitively penetrates (has a causal impact on the content of) their perceptual seemings. She argues that in such cases the agent's beliefs are unjustified, and perhaps even circular. I argue that analogous cases to those given by Siegel can be given for testimonial belief, and that this causes a problem for particular forms of testimonial anti-reductionism (and some conceavable brands of reductionism). Views of testimonial justification are taxonomized according to the particular default rights which are postulated, and the different stances which might be taken on the epistemic rights of agent in cases of misunderstanding. Moreover, two forms of cognitively penetrated linguistic understanding are identified - one in which one's speech perception is penetrated, and one in which irrational background states have an impact on the assignment of meaning to words. It is argued that cases of cognitively penetrated linguistic understanding cause a problem for views of testimonial justification which postulate a default right for agents to trust their own understanding.

Chapter Two: Testimonial Knowledge Without Knowledge of What is Said.

0. Introduction.

In this chapter I discuss the following question: what epistemic relation must audiences bear to the contents of assertions (and the fact that those contents have been asserted) in order to gain testimonial knowledge from those assertions?   I briefly explain why epistemologists should focus more attention on this question, before providing three counterexamples to the most intuitive answer: that audiences must know that a speaker has asserted *p* in order to gain testimonial knowledge that *p*. I close by explaining how the argument generalises and can be made to work on different assumptions about the conditions for knowledge, and the conditions under which a proposition is asserted.

1. Audiences and Asserted Contents.

Standard formulations of testimonial knowledge rarely make any mention of the relation audiences must bear to an asserted content (or the fact that it has been asserted) in order to gain testimonial knowledge. Rather, focus is usually on the relation the audience must bear to the speaker (i.e. must audiences have reason to trust speakers?), and the relation the speaker must bear to the asserted content. This is perhaps due to the fact that content recovery is usually taken to be unproblematic by epistemologists. As Sanford Goldberg puts it:

'...so prevalent is the assumption that the comprehension dimension is unproblematic - that hearers reliably recover the propositions attested to - that there is virtually no discussion of the comprehension processes in the epistemological literature' Goldberg (2007), 54.

However, as noted in the introduction, in the philosophy of language content recovery is not seen as epistemically unproblematic. Theorists concerned with context-sensitivity have long recognised that audiences sometimes struggle to recover the precise proposition intended by the speaker. Anne Bezuidenhout lists the following factors as contributing to audiences' interpretations of context-sensitive utterances. All of these factors leave space for misinterpretation:

'(i) Knowledge that has already been activated from the prior discourse context (if any).

(ii) Knowledge that is available based on who one's conversational partner is and on what community memberships one shares with that person.

(iii) Knowledge that is available through observation of the mutual perceptual environment.

(iv) Any stereotypical knowledge or scripts or frames that are associatively triggered by accessing the semantic potential of any of the expressions currently being used.

(v) Knowledge of the purposes and abilities of one's conversational partner (e.g. whether the person is being deceitful or sincere, whether the person tends to verbosity or is a person of few words etc.).

(vi)   Knowledge one has of the general principles governing conversational exchanges (perhaps including Grice's conversational maxims, culturally specific norms of politeness, etc.)'. Bezuidenhout (2002), 117.

In chapter one we also saw psychological evidence that verbal communication is less reliable than we often assume. The controversy over communication with context sensitive terms illustrates that, despite epistemologists' assumptions to the contrary, it is not clear that comprehension is unproblematic. Indeed, it may even be that the common assumption among epistemologists that the recovery of propositional contents is generally reliable is partly due to our general tendency to over estimate our reliability as communicators (discussed by Savitsky, Keysar, Epley, Carter, and Swanson (2010)).   Thus the question of the audience's relation to what is said is more pressing than has previously been recognised by epistemologists (we will return to these epistemic problems with content recovery in more detail in chapters three and four).

Moreover, it is important to have some idea of the relation audiences must bear to asserted contents when investigating the epistemic impact of the processes underlying interpretation. This is because epistemic problems arising from these processes will often prevent audiences from bearing more epistemically demanding relations to asserted contents. Thus, if testimonial knowledge requires a very demanding relation to what is said then problems with the processes underlying our communicative exchanges will often prevent us from gaining testimonial knowledge. However, if audiences are not required to bear an epistemically demanding relation to what is said we might find that the sorts of problems with content recovery discussed in this thesis do not necessarily block testimonial knowledge (of course, we may still find that they often do, as a matter of contingent fact, block testimonial knowledge, as I will argue in chapter three).

Finally, it is important to know what relation audiences must bear to what is said (and the fact that it has been said) in order to assess the a priori/a posteriori status of testimonial knowledge in general. For example, if knowledge of what is said is required for testimonial knowledge, and knowledge of what is said is always a posteriori, then presumably all testimonial knowledge is a posterori. If some lesser relation is all that is required we would have to ask about the potential a priori status of this lesser relation. For example, if mere propositional justification that *p* has been said (or presented as true) is required then we must ask whether or not such propositional justification can be a priori. I will not be perusing this question further here, but it does illustrate another sense in which the question is important.

Perhaps there is a simple answer to this question: Hearers must know (or be in a position to know) that the speaker has said that *p* in order to gain testimonial knowledge that *p*[16]. After all, if the audience doesn't even know that the speaker has said that *p* then why worry about whether or not the speaker is trustworthy or reliable with respect to *p*? This seems like a highly intuitive requirement on testimonial knowledge. Moreover, when we are told that *p,* and we are asked why we believe that *p*, we will usually state '*S* said that *p*' as our reason. If the fact that a speaker said that *p* is part of our reason for believing that *p* then surely knowledge that the speaker said that *p* is necessary for testimonial knowledge that *p*. Anna-Sara Malmgren (2006) employs this thought in her argument against Tyler Burge's claim that there is a priori testimonial knowledge. She claims that knowledge of what is said is a posteriori, and that knowledge of what is said plays an epistemic

---

[16]   This answer is already endorsed by Fricker (1994, 1995, and 2003), Christensen and Kornblith (1997), Malmgren (2006), and Goldberg (2007).

role in the acquisition of testimonial knowledge, meaning that a posteriori knowledge plays an epistemic role in testimonial knowledge. If Malmgren is correct then there cannot be a priori testimonial knowledge (we will return to Burge's arguments for a priori testimonial knowledge in the final chapter where I provide a new response to his argument). Malmgren sums up the thought as follows:

> 'However, the following consideration gives us a prima facie reason to think that it plays an epistemic role: suppose John tells me that it is raining, and that I thereby come to know that it is raining—that is, suppose that I gain knowledge by (John's) testimony that it is raining. If you asked me *how* I know that it is raining, then presumably part of my (pretheoretical) answer would be: "John told me," "John said so," or "John said that it is raining." Here is a natural thought about *what I am doing* in giving this answer: I am citing part of my (epistemic) *reason* for believing that it is raining, part of what makes me *warranted* in believing that this is the case. What my answer brings out is that part of my reason for believing that it is raining is that John said so. But that is just to say that my warranted (or knowledgeable) belief about what John said plays an epistemic role in the formation of my knowledge that it is raining.' Malmgren, (2006), 225.

Furthermore, it might be worried that if an audience doesn't know that *p* has been asserted then they would have believed *p* even if *p* had not been asserted (and thus in cases in which *p* is false)[17]. A knowledge requirement is also likely to appeal to the accessibility internalist (for example Bonjour 1980). Such theorists hold that in order to know that *p* one must be able to come to know the basis for one's knowledge that *p*. It is commonly thought that the fact that the a speaker has testified that *p* is at least part of the basis for one's testimonial knowledge that *p*. Thus, in order to gain testimonial knowledge that *p* one must at least be in a position to know that the speaker has said that *p*.

Despite its intuitive plausibility this answer is incorrect. Audiences needn't be in a position to know that a speaker has asserted that *p* in order to gain testimonial knowledge that *p*. This is a relief, since the arguments of the following two chapters suggest that we very often lack strict knowledge of what is said, which means that we would in principle be blocked from gaining testimonial

---

[17] Goldberg (2007) pursues this line of argument, and I will consider it in a little more detail in chapter three.

knowledge. I will argue that testimonial knowledge does not require knowledge of what is said by providing three cases of apparent testimonial knowledge of *p* in which the audience does not know that the speaker has said that *p*.

2. Against the Knowledge Requirement.

The main argument of this chapter rests on two primary assumptions. Firstly I assume a roughly Gricean intention based view of communication. That is, I assume that in order for a speaker to say that *p* they must intend to say that *p* (at least for utterances containing context-sensitive terms such as demonstratives and definite descriptions). The second assumption depends on the first: if an audience member is not in a position to know whether the speaker intended to say that *p* then the audience is not in a position to know that the speaker said that *p*. If, for all the audience knows, the speaker did not intend to say that *p,* then since the speaker's intention to say that *p* is necessary for their having said that *p*, for all the audience knows the speaker did not say that *p*. For the majority of the chapter I also assume that safety is a necessary condition on knowledge. I adopt these assumptions because the Gricean view of 'what is said' is the dominant view in the philosophy of language, and the safety principle on knowledge is still the dominant anti-luck condition in epistemology. Moreover I am sympathetic to both views. However, as as I note in 2.1.2 and the conclusion, these assumptions are not essential. This is important since it is easy to worry that I am attacking a straw man. That is, it is easy to read the cases presented here and feel that very few people are committed to the view that testimonial knowledge requires knowledge of what is said where 'what is said' is spelled out in the terms presented here. This may be correct (although the Gricean view certainly seems to be the default position on 'what is said' for many). However, as long as one employs a notion of 'what is said' in the conditions one gives for testimonial knowledge one will need to give an account of the conditions under which a proposition counts as 'what is said'. As I explain in the conclusion cases similar to the ones given here will be available for any set of conditions which leave space for the audience to be mistaken about whether or not *p* has been said.

I will present three cases in which the audience appears to gain testimonial knowledge that *p* despite not being in a position to know that the speaker intended to communicate that *p*. If the two aforementioned assumptions are correct then these will also be cases in which the audience gains testimonial knowledge that *p* without knowledge that the speaker said that *p*. I consider three cases primarily for the sake of clarity. Considering several cases allows us to identify a common structure

for counter examples to the knowledge of what is said requirement. This allows us to see how we might generate further counter examples based different assumptions about the conditions for what is said, and the conditions for knowledge. Moreover, intuitions about cases are notoriously unreliable, especially intuitions about cases involving far fetched sci-fi scenarios or cases with very complex set-ups. Thus it is worth considering several cases in order to ensure that we are not led astray by some misleading feature of one case. Finally, each case gives rise to a different intuitive line of response. These responses are considered and rejected. However, the reader might not find all of the counter-responses convincing. Thankfully, they need only find one of the three counter-responses convincing in order for the project of this chapter to be successful.

2.1.1. Case One.

MAD SCIENTIST[18],[19]: The philosophical mad scientist is at it again. His victim is Sally, a car enthusiast. This time, instead of envating his victim, he has implanted a special chip in her brain. This chip causes her to sometimes say 'that is a fuel efficient car' when in the presence of fuel efficient cars, but only when she does not intend to do so. It works as follows: whenever Sally is in the presence of a fuel efficient car it turns on and randomly selects one of two values. If it selects value 1 it switches off and becomes inactive again. However, if it selects value 2 it has the following effect: If Sally doesn't intend, and doesn't gain the intention, to comment on the fuel efficiency of the car, then it forces her to utter the sentence 'that is a fuel efficient car'.

One day Sally and Matt are walking through the city when Sally sees a particularly fuel efficient car. She considers commenting on its fuel efficiency, but hesitates because she doesn't know if Matt has any interest in cars. She decides on a whim to

---

[18] Incidentally, this case also constitutes a counter example to Jeff King's (forthcoming) claim that in order to say that *p* a speaker must intend to say that *p* and the audience must be in a position to know that the speaker intends to say that *p*. This is because the speaker in this case clearly says that *p* even though the audience member is not in a position to know that the speaker intended to say that *p*.

[19] The following cases are similar to those provided by Federico Luzzi (2010) as counterexamples to counter-closure (the view that in order to gain knowledge that *p* by inference from *q* one must know *q*). Indeed, if one held a view of testimony on which audiences make a sub-personal inferences from beliefs like 'S said that *p,* and S is trustworthy' to 'I should believe that p', then one would be forced to reject counter closure on the basis of the cases presented in this chapter.

just go for it, and says 'that is a fuel efficient car!'. On the basis of Sally's assertion Matt forms the true belief that the car is fuel efficient. What neither Sally nor Matt know is that this was a case in which the chip selected value 2, so Sally would have uttered 'that car is fuel efficient' even if she did not intend to.

I submit that Matt gains testimonial knowledge that the car is fuel efficient. His belief about the fuel efficiency of the car was formed on the basis of Sally's testimony, and it is safe and sensitive (he forms the same belief in all nearby worlds, it is true in all nearby worlds, and in the closest words where it is false neither Sally's intentions nor the chip cause her to utter the sentence 'that is a fuel efficient car'). His belief is causally related to its truthmaker, and Sally undertook all the usual commitments one undertakes when one provides testimony. Indeed, as emphasised by Luzzi (2010), if cases like this don't count as cases of knowledge then they they constitute interesting new forms of Gettier cases not answerable in any of the standard ways.

However, Matt does not know (nor is he in a position to know) that Sally intended to communicate that the car was fuel efficient. This is because his belief that she intended to communicate that the car was fuel efficient was neither safe nor sensitive[20]. In the closest worlds in which Sally doesn't have the intention to say that the car is fuel efficient he still forms the belief that she does intend to say that it is fuel efficient, because the chip makes her say it. Moreover, since Sally was very hesitant about saying that the car was fuel efficient and only decided to do so on a random whim there are many nearby worlds in which she doesn't form the intention to comment on the fuel efficiency of the car. Thus, if assumptions 1 and 2 are correct then Matt gains testimonial knowledge that the car is fuel efficient despite not knowing that Sally said that the car is fuel efficient.

2.1.2. Response to Case One.

It might be argued that this case, rather than being a a case of testimonial knowledge without knowledge of what is said, is a counter example to the safety principle on knowledge. That is, it might be argued that although Matt's belief about Sally's intention is unsafe he nonetheless knows

---

[20] One might worry that audiences do not generally form beliefs about speaker intentions when they form testimonial beliefs. This worry would be consistent with the view that audiences must be in a position to know the speaker's intentions in order to gain testimonial knowledge. However, these cases illustrate that the audience need not be in a position to know the speaker's communicative intentions, since any belief they may potentially form about the speaker's communicative intentions will fail to be safe or sensitive.

that Sally intended to communicate that the car was fuel efficient. Indeed, this case does bear some resemblance to a series of cases which have been employed as counterexamples to the safety condition on knowledge (see Neta and Rohrbaugh (2004) and Comesana (2005))[21]. For example, Neta and Rohrbaugh present the following case:

> EXPERIMENT: 'I am participating in a psychological experiment, in which I am to report the number of flashes I recall being shown. Before being shown the stimuli, I consume a glass of liquid at the request of the experimenter. Unbeknownst to either of us, I have been randomly assigned to the control group, and the glass contains ordinary orange juice. Other experimental groups receive juice mixed with one of a variety of chemicals which hinder the functioning of memory without a detectable phenomenological difference. I am shown seven flashes and judge, truly and knowingly, that I have been shown seven flashes. Had I been a member of one of the experimental groups to which I was almost assigned, I would have been shown only six flashes but still believed that I had been shown seven flashes due to the effects of the drug. It seems that in the actual case I know that the number of flashes is seven despite the envisaged possibility of my being wrong. And yet these possibilities are as similar in other respects as they would have to be for the experiment to be well designed and properly executed.' Neta and Rohrbaugh, (2004), 400.

I find this case (and others like it) unconvincing, since I find myself unable to shake the feeling that the agent in such cases does not genuinely know[22]. However, the issues at play in this debate are subtle, and I do not want to rest my case on my own (no doubt theory laden) intuitions. Luckily I don't have to. Neta and Rohrbaugh present a modified version of their case (analogous to

---

[21] There are further examples provided in Kelp (2009), Baumann (2012), and Bogardus (2014). However, these cases bear less resemblance to the cases presented here.

[22] I feel differently about Comesana's case. In Comesana's case Juan is planning to dress up as Michael for a Halloween party, but decides against it at the last minute. Judy, has been employed to give directions to the party. She will tell everyone it is in the same location, but if she sees Michael (or someone she thinks is Michael) she will ring the organiser and the party location will be moved. So although someone looking like Michael would receive the same testimony as everyone else, Judy would arrange for that testimony to be false. Since Juan almost dressed up as Michael there is a sense in which he almost received false testimony. However, it strikes me that Juan's belief in this case is not unsafe. The world at which he receives the false testimony is just too different from actuality. Indeed, it is not even clear to me that the worlds in which Juan receives the false testimony are worlds in which the same belief forming method is employed. This is because when we form testimonial beliefs we rely on the speaker to be honest to *us* - the audience. When Juan comes as himself he relies on Judy to be honest to him - to Juan. However, if he dresses up as Michael he relies on Sally to be honest to Michael. Indeed, we could even imagine that Michael knows of the plan to keep him from the party, and decides to dress up as Juan in order to receive truthful testimony as to the party's location. In this case Michael is clearly and intentionally employing a different method than he would if he approached Judy undisguised. Thus a different belief forming method is employed in each case. Thus, since safety is always relativised to a method, Juan's belief is safe (for views which place a special emphasis on this interpersonal relation between the speaker and the hearer see Moran (2005a, 2005b), Hinchman (2005), and Fricker (2006)).

Goldman's (1976) Fake Barns case) in which it seems that the agent clearly does not know that they have been shown seven flashes. The case is as follows:

> EXPERIMENT*: 'I am taking part in a long series of psychological experiments, in each of which I am to report the number of flashes I recall being shown to me after ingesting a glass of liquid. In this one case, I have been assigned to the control group and the liquid is ordinary orange juice. I am shown seven flashes and judge, truly, that I have been shown seven flashes. In some of the other trials in which I have participated, I have been assigned to an experimental group in which the liquid also contains a drug which interferes with memory, and the beliefs I formed on those trials were false'. Neta and Rohrbaugh, (2004), 402.

We can modify the mad scientist case in a similar way in order to secure the intuition that Matt does not know that Sally intended to comment on the car's fuel efficiency. Imagine that Sally and Matt have spent the afternoon together, and Sally has commented on the fuel efficiency of many cars. In some of these cases she has been caused to vocalise the sentence 'that is a fuel efficient car' as a result of the chip's interference, in other cases she did so of her own volition. We can even imagine that Sally's hesitation to comment on the fuel efficiency of the car was due to her noticing Matt's seeming lack of interest in her previous comments on fuel efficiency[23]. In this version of the case Matt will have formed many false beliefs about Sally's communicative intentions, and he will have been lucky in those cases where the beliefs he formed were true. This is clearly a situation in which his faculties for assigning communicative intentions are not reliable. Thus it seems highly implausible to claim that he knows that Sally intended to comment on the car's fuel efficiency.

Nonetheless, Matt's actual belief about the fuel efficiency of the car is still safe and sensitive. After all, the chip never makes Sally utter 'that is a fuel efficient car' when she is not in the presence of a fuel efficient car. Additionally, his belief is still caused by its truthmaker, and the causal chain which actually lead to his belief was not deviant in any way. Finally, Sally undertook all the usual commitments when she asserted. Thus, it seems that Matt still gained testimonial knowledge, despite the fact that he did not know Sally's communicative intentions.

A second worry about this case may be that when the chip in Sally's brain causes her to say 'that is a

---

[23] We can imagine that the chip retroactively implants in Sally the belief that she intended to comment on the car's fuel efficiency when she is forced by the chip to do so.

fuel efficient car' she is not really testifying. Rather, in such cases she is acting as an instrument. If this were the case then Matt's belief would not be formed via the same method in each case. This is not a problem. In all the worlds in which Matt does form a belief about the fuel efficiency of the car on the basis of Sally's actual intentional utterance his belief is still true. So his belief about the fuel efficiency of the car is still safe. Matt might be thought to be lucky in one sense - he was lucky that his belief was testimonial. But he was not lucky that it was true. Moreover, although he might be thought to gain knowledge about the car via a different method in the worlds in which Sally never forms an intention to assert that the car is fuel efficient, his belief about Sally's actual intentions is formed via the same method, for his evidence regarding Sally's intentions is the same in each case (his evidence being the fact that she opened her mouth and the words 'that is a fuel efficient car' came out). So the chip worlds are relevant to the assessment of the safety of his belief about Sally's intentions.

2.2.1. Case Two.

I take MAD SCIENTIST to straightforwardly establish that knowledge of what is said is not required for testimonial knowledge. However, it is a rather far fetched sci-fi case. Perhaps we shouldn't trust such far fetched cases. Thankfully a similar case can be given which does not appeal to mad scientists and their bizarre mind control devices:

> COMPETING INTENTIONS: In this case Sally does not have a chip in her head, rather she just impulsively talks about cars. She almost always comments when she sees a fuel efficient car. Sally and Matt are walking through the city when Sally notices two cars which she knows to be fuel efficient. One of the cars (car A) is partially obscured by a wall, and has no stand out features. The other car (car B) is in full view, and has a very striking design. It is inevitable that Sally will comment on the fuel efficiency of one of the cars. Suppose that both cars have the same owner, he takes these two cars everywhere he goes. If either car were absent then both cars would be absent. Suppose further that Sally takes both cars to be highly salient, because although car A looks rather plain it is actually very rare, and she falsely takes this to be common knowledge. Finally, suppose that Car A and car B are different editions of the same car but with different body work. Car A is a limited edition of car B. Thus they both have identical engines, and everything else related to fuel efficiency. As a result, if car B were inefficient then car A would be inefficient too (and vice versa).

Sally is torn as to which car to comment on, but in the end (and in a random manner as in the first case) decides to comment on car B, stating 'that is a fuel efficient car'. Car B is the only car which is salient to Matt, so he correctly takes Sally to be commenting on car B. However, Sally is a rather poor communicator: if she had decided to comment on car A (which she very nearly did), she would have done so by uttering precisely the same sentence. Since the cases are identical from Matt's perspective he still would have taken her to be talking about car B, and would still have gained the true belief that car B is fuel efficient.

This case also seems to be one in which Matt gains knowledge. Once again his belief is formed on the basis of Sally's testimony, it is true, and Sally undertook a commitment to the proposition Matt came to believe. Additionally his belief seems to be safe - he forms the same belief in all nearby worlds, and it is true in all these worlds, and sensitive - in the closest worlds at which his belief is false he would not have formed it, since the closest worlds in which car B is inefficient are worlds in which car A is inefficient as well, meaning that Sally would never have uttered the words 'that car is fuel efficient'. However, Matt's belief that Sally intended to say that car B is fuel efficient, although true, does not constitute knowledge due to failing safety and sensitivity. Sally very nearly intended to say that car A was fuel efficient, but due to her low competence as a communicator Matt would still have taken her to be saying that car B was fuel efficient.

2.2.2. Response to Case Two.

One potential worry about this case is that Sally offers different testimony in the worlds in which she intends to comment on car A and the worlds in which she intends to comment on car B. I have maintained that Matt's belief about Sally's intentions is unsafe because he would have formed the same (false) beliefs in worlds where Sally decided to comment on car A. However, safety is standardly relativised to a method. That is, the important thing is not whether there are nearby worlds where Matt forms a false belief, but rather whether there are nearby worlds in which Matt forms a false belief by employing the same method. If Sally offers different testimony in each case then surely Matt's belief is not formed via the same method.

This objection fails for two reasons. Firstly, Matt's unsafe belief concerns Sally's communicative intentions, whereas Sally's testimony concerns the fuel efficiency of a car. The mechanisms Matt would use to reach a judgement about Sally's communicative intentions are identical in each case. Matt has precisely the same evidence in each case, the only difference is that in the one case his belief about Sally's intentions is correct, and in the other case it is incorrect. Matt's belief about Sally's communicative intentions is not a testimonial belief, so the claim that she offers different testimony in each case is irrelevant (this response, of course, is the same as the response to the second objection in section 2.1.2).

Secondly, even if Matt's belief regarding Sally's communicative intentions were testimonial there is reason to doubt that this objection works. It seems that when considering the safety of testimonial beliefs the speaker's communicative intentions should not be built into the method component of our safety assessment. There are cases in which the speaker has very different communicative intentions, yet it seems clear that the same method is employed by the audience when forming their testimonial belief. This is not to say that the speaker's communicative intentions are epistemically irrelevant, but there are many roles such intentions can play which don't require them to be built into the method component of our safety assessments for testimonial knowledge. Consider the following cases:

> ANCIENT SCHOLAR: Archaeologists have just uncovered a manuscript detailing the rise and fall of ancient civilisation x. They have extremely strong evidence that it was written by Heromomouse, an ancient scholar belonging to a different civilisation – ancient civilisation y. They know Heromomouse to be extremely reliable. They read the parchment and learn many facts about ancient civilisation x.

> Version 1: Heromomouse wrote the manuscript in order to educate the masses about the history of ancient civilisation x, and he intended the manuscript to be kept in the great public library.

> Version 2: Heromomouse intended the manuscript as a gift for the king's private library, and he intended it to be read only by the king. Moreover, he does not expect (or even intend) the king to believe a word of it. Heromomouse knows that the king is far too pig headed to believe that there was ever a civilisation greater than his own. Nonetheless, Heromomouse ensures that the document is as accurate as possible out of a patriotic desire for his king to have the greatest private library in the land.

Version 3: This version of the case resembles the second in all respects except that Heromomouse knows that the king will have him beheaded at any suggestion that there was a civilisation greater than his own. However, he also knows that the king is so delusional that he will interpret Heromomouse's chronicling of civilisation x as a metaphor for the rise and fall of his own civilisation. Heromomouse knows that no sane person would read the manuscript this way, but this doesn't matter since the manuscript is for the king's private library. He still wishes for the king to have a great library, thus he still strives for accuracy. However, he intends to be interpreted as expressing an entirely different message.

In each version of this case Heromomouse's communicative intentions are rather different. In the first case he is openly offering testimony regarding the rise and fall of civilisation x. In the second case he does not intend his manuscript to be read by anyone other than the king, and he does not intend to be believed. In the final case he does not even intend to be interpreted the same way, he intends to communicate something completely different to the other two cases. If we wish to claim that the audience employs a different belief forming method in cases where the speaker's communicative intentions are different then we must claim that the archaeologists employ different methods in each of these cases. Yet, this is surely false. In each case the archaeologists employ the same method to learn about civilisation x. Thus, the fact that the speaker has different communicative intentions in each case does not entail that the method employed by the audience in each case is different. So there are two reasons to reject the 'different methods' response to COMPETING INTENTIONS. Firstly Matt's knowledge of Sally's intentions is not testimonial, so the difference in communicative intentions is irrelevant. Secondly, although they no doubt do play an important epistemic role, communicative intentions should not be built into the method component of safety judgements about testimonial knowledge.

2.3.1. Case Three.

COMPETING INTENTIONS is less far fetched than MAD SCIENTIST. However, it is still rather complex. Thus, I imagine that the dedicated proponent of the knowledge condition might well remain unconvinced[24]. Additionally, the question of the audience's knowledge of what is said is of

---

[24] Another strategy my opponent could take would be to question whether the knowledge gained in this case is genuinely testimonial, since it depends on so many outside factors (thanks to Brian Weatherson for this point). Here is how I think we should conceive of testimonial knowledge: testimonial knowledge is a testimonial belief which counts as knowledge. A testimonial belief is a belief formed on the basis of testimony, which means something like the following: Someone has said that *p*, or made an utterance with the content *p*, or intended to say that *p* and had their intention recognized, or said something sufficiently similar to *p* and the audience reasonably recovered *p* etc.

interest due to worries raised in the philosophy of language about the successful recovery of the speaker's intended meaning. Yet, neither of the cases so outlined so far bears any similarity to the problematic cases discussed in the philosophy of language. Thus it is worth considering a final case which bears directly on problems of context-sensitivity[25].

The reader will recall the problem from section one. The problem is that in cases of context-sensitivity there will often be several similar values which can be assigned to a context-sensitive term, each determining a different proposition[26]. The audience's evidence will often be insufficient to grant them knowledge that the speaker intended one of these propositions rather than another. However, as the final case indicates, these epistemic difficulties do not always rule out testimonial knowledge

As noted in the introduction, Wettstein (1981) provides the following argument against the Russellean view of definite descriptions: in most attributive uses of a definite description the explicit content underdetermines the referent. That is, it is not sufficient to determine a unique referent. For example, take the sentence 'the statue is broken' - the world contains many statues, thus 'the statue' is not enough by itself to select a unique statue. Those words could be used in different contexts to identify very different statues. The description must be supplemented in order to designate a unique statue. However, there will be many potential supplementations which would uniquely identify the intended statue. For example the description might be supplemented as 'the statue belonging to Alex', or 'the statue of a frog'. Wettstein argues that this will make the proposition expressed indeterminate on the Russellean view. He takes this to be an unacceptable consequence and provides an alternative view of definite descriptions on which the meaning is determinate. However, the problem seems to arise for a wide array of context sensitive terms, not just descriptions. Similar cases can be given involving quantifier domain restriction, non-sentential assertion, and a whole host of other context sensitive terms. Thus Wettstein's response seems unsatisfactory, for it is unlikely that a unified account can be given which secures determinacy for

---

and on the basis of this fact the audience came to believe *p*. On this conception of testimonial knowledge Matt's knowledge is genuinely testimonial.

[25] The cases discussed so far have also been cases of context sensitivity. However, they are cases of reference assignment which seems less epistemically problematic than many other forms of context sensitivity. In cases of reference assignment there are usually fewer candidate meanings for the audience to choose between. In the cases which are generally seen to be problematic there are a large number of similar meanings for the audience to choose between. It is seen as mysterious how the audience could select the correct meaning from amongst the many available candidates.

[26] This will not generally be the case when context merely supplies a referent (as is the case with, for example, demonstratives or referential uses of definite descriptions). However, it will be the case when context must supply a more complex value such as a property, comparison class, modal base, ordering relation on worlds, standard of taste etc, and in cases of loose talk or non-sentential assertion.

all context sensitive terms. Thus a better response seems to be to accept that such indeterminacy is a feature of context sensitive terms. This can lead to problems concernig knowledge of what is said however. Consider the following case (adapted from Donnellan (1966)):

> MURDERER: Sally and Matt are investigating a chain of murders they know to be related. Sally is standing over the mangled remains of Frank - the latest victim. Frank has clearly been murdered in a rather brutal way, and the scene has been covered in insane etchings and other clear indicators that the murderer is insane. Additionally, both Sally and Matt know that Frank was murdered by the murderer they are investigating (that is, the individual who committed the previous murders). Sally phones Matt and says 'the murderer is clearly insane'. The description used here is incomplete, and could be rendered complete in a number of different ways. For example, Sally could have intended the proposition 'the murderer (of Frank) is clearly insane', or 'the murderer (who we are investigating) is clearly insane'. Matt recovers the proposition 'the murderer (who we are investigating) is clearly insane' (which happened to be the proposition intended by Sally), and comes to believe it on the basis of Sally's testimony.

Once again I think that it is clear that Matt gains testimonial knowledge in this case. The belief he forms is true, safe, sensitive, and caused by its truth maker. Additionally, Sally undertook all the usual commitments to the proposition that Matt comes to believe. However, since as far as Matt can tell Sally might have intended the other (very similar) proposition, Matt does not know that Sally intended the proposition he recovered.

2.3.2. Response to Case Three.

There are several possible responses to this case, however each response fails. Firstly one might attempt to maintain that this case involved a referential use of the definite description. If this were the case then there would not be multiple competing candidate meanings, there would be one single proposition expressed. This response does not seem promising since neither Sally nor Matt know who the murderer is, they merely know that there is a single murderer responsible for all the murders they are investigating. Thus, it does not appear that Matt or Sally have the resources with which to singularly refer to the murderer. Nonetheless, this may just push the proponent of the knowledge of what is said condition to adopt a very liberal view of the conditions required for

singular reference. Thankfully we needn't rest on the claim that this case involved an attributive use of the definite description, for a directly analogous case can be given with quantifiers. Consider the following:

> MURDERER*: In this case Sally and Matt are once again investigating a chain of murders they know to be related. Every victim is a member of 'The secret Society of Evil', and the same insane etchings have been found at each crime scene. However, in this case there is significant dna evidence which suggests that (and which leads Matt and Sally to believe that) different murderers committed each crime. This time Sally phones Matt and says 'Every murderer is clearly insane'. Clearly Sally is not saying that every murderer in the world is insane, rather she is saying that a certain group of murderers are insane. Thus, in this case a property must be supplied to restrict the quantifier. However, there are several equivalent properties which Sally could intend. For example, she could intend 'Every murderer (we are investigating) is clearly insane', or 'Every murderer (of a secret society of evil member) is clearly insane'. Matt recovers 'Every murderer (we are investigating) is clearly insane', and as it happens this is the proposition Sally intended.

Once again it seems clear that Matt gains testimonial knowledge. However, it also seems clear that he did not know Sally"s communicative intention, for she could, for all he knows, easily have intended to communicate 'Every murderer (of a secret society of evil member) is clearly insane'. However, this time the referential use response is not available[27].

A second response might be to hold that in a case like this Sally really asserted both propositions. One might hold that in typical cases a speaker actually asserts many different propositions (for a view like this see Cappelen and Lepore (2006)). This seems intuitive, since it certainly seems strange to imagine Sally specifically intending one of the above readings and not the other. They seem to come together. I do not believe that this undermines the case however. The case is a simplification, and we can see that when the case is rendered more realistic the problem remains. It is unlikely that Sally intended or asserted only one proposition. It is more likely that there is a set

---

[27] One might worry that context does not supply a property to restrict the quantifier, but rather supplies a the domain restriction in the form of a set. This would be a version of the referentialism response for quantifier domain restriction. Indeed, through much of Stanley and Szabó's seminal (2000) paper on quantifier domain restriction they speak as if this is their view. However, as they explain, this is a simplification. They ask us to consider the following case: John buys 70 bottles of beer every time he goes to the supermarket. This time there are only 70 bottles of beer on the shelf, and so he buys every bottle of beer. However, someone could truly utter 'If there were a few more bottles on the shelf John would not have brought every bottle'. If quantifier domain restriction merely provided a set this sentence could not be truly uttered, because 'every beer' would pick out the same set of beers in worlds where there were more than 70 beers on the shelf (Stanley and Szabó)2000), p 252) .

of propositions consistent with Sally's communicative intentions. However, just as there seemed to be a problem with Matt recovering the precise proposition Sally intended there seems to be a problem with him recovering the precise set of propositions which are consistent with Sally's intentions, or with him recovering only propositions which are in this set. We can conceive of cases in which, for all Matt knows, the proposition he recovers might not be in the set of propositions consistent with Sally's intentions. For example, imagine that Sally and Matt both have access to a huge amount of information about the murderer, and this causes them to think about the murderer in different terms at different times. Matt has spent all day looking at the connections between the murders and several gangs, and has been thinking of the murderer as 'the gang affiliated murderer'. He has not been thinking about the murderer in terms of the many financial crimes which the murderer has committed. Although he knows the murderer has committed a set of financial crimes he often has to do some cognitive work to recover this information. Sally on the other hand has spent all day thinking about the financial ties in the case. However, just before she phones Matt she sees a gang sign on the victim's arm and this raises the gang affiliations back to salience. As a result the resolutions of the description which relate to gang affiliation do happen to fall under her intention, and thus the proposition Matt recovers does fall under Sally's intention. However, it easily could have failed to do so. And in this situation Matt still would resolved the context sensitivity in the same way. So, for all Matt knows the proposition he recovers is not in the set of propositions consistent with Sally's communicative intentions. Yet, it still seems he gains testimonial knowledge.

3. Conclusion.

It has been argued that despite its intuitive appeal, the claim that audiences must know what has been said in order to gain testimonial knowledge is false. The argument presented here assumed that a speaker must intend to say that *p* in order to say that *p* (at least in cases involving context sensitivity). I also assumed a safety condition on knowledge through most of the chapter. I adopted these assumptions because the Gricean view of 'what is said' is the dominant view in the philosophy of language, and the safety principle on knowledge is still the dominant anti-luck condition in epistemology. Moreover I am sympathetic to both views. However, we are now in a position to see that these assumptions were not essential. The cases I presented all have the following features: a speaker asserts a proposition *p,* the speaker meets all the normal requirements for the transmission of knowledge (e.g. the speaker knows that *p*, and commits to *p* etc.), the audience recovers *p*

through the usual means, is justified in believing that *p* has been asserted, and forms a belief that *p*. Moreover, the audience's belief that *p* meets the usual conditions on knowledge (e.g. safety, sensitivity, being caused by its truthmaker, local reliability etc.). However, there is some necessary condition C for a speaker's saying that *p* such that the audience is not in a position to know that C is met.

In the cases I presented C was a speaker intention requirement. However, as long as it is possible for a speaker to say that *p* without the audience thereby coming to know that *p* has been asserted it seems that it will likeley be possible to generate similar cases with different C conditions. For example, if we confined ourselves to cases involving demonstratives and adopted a view on which a demonstration is required to pick out the referent of a demonstrative (for example, Reimer (1991)), then we could produce cases in which an audience is not in a position to know what object the demonstration picks out, but in which the testimonial belief they form nevertheless meets the conditions on knowledge. Alternatively we might suppose that very small differences in the patterns of use give rise to subtle changes in meaning (see Dorr and Hawthorne (2014) for a discussion of the consequences of such views). Then we might consider cases in which the audience is not in a position to know that the pattern of use is such as to determine the meaning they assign, but in which they still seem to meet the conditions for knowledge with respect to the object of their testimonial belief. These will all be cases in which any belief an audience might potentially form regarding condition C will fail to meet some condition on knowledge (be it safety, sensitivity, local reliability, some causal condition etc.), but in which some feature of the environment guarantees that their belief in the proposition testified to does meet all the conditions for knowledge. This is important, since it is easy to worry that the argument presented in this chapter attacks a straw man. That is, it is easy to worry that no epistemologists of testimony explicitly endorse the view that testimonial knowledge requires knowledge of what is said where *p* must be intended in order to count as 'said'. This might be true, but as long as epistemologists of testimony appeal to a notion of 'what is said' they will need some account of the conditions under which *p* counts as being said. And an argument along the lines of the one presented here will be available for most sets of conditions which leave open the option of *p* being said without the audience to be in a position to know that *p* has been said. Moreover, the most natural fall back position would be to move away from the notion of 'what is said' and formulate theories of testimony in terms of the speaker's intended meaning. However, such an approach would also face the arguments presented here, since they show that knowledge of the speaker's communicative intentions is not required for testimonial knowledge.

Thus, attempts to solve the problem by adopting a different metasemantics for context sensitivity, or by rejecting the safety condition on knowledge, seem unpromising[28].   One might reject the assumption that a speaker can say that *p* without the audience being in a position to know that they have said it (a similar condition is endorsed by King (2014, forthcoming)).   However, in order to do so one would have to not only deny that testimonial knowledge was gained in any of the cases outlined in this chapter, but also deny that anything was said. This seems like a radical move. So, it seems the best response is to embrace the notion that knowledge of what is said is not required for testimonial knowledge. It is not clear what alternative relation (if any) an agent must bear to the object of testimony in order to gain testimonial knowledge. However, in order for a proposal to be immune to the sorts of counterexamples presented here it seems the best strategy will be to search for a relation which is entailed by the collective presence of the other preconditions for testimonial knowledge (including the speaker's having said that *p*). Otherwise there will be space to present a case in which the other conditions for testimonial knowledge are met (including the causal, anti-luck, and reliability conditions on knowledge) in which the condition does not apply. In such a case we will need strong reasons to deny that testimonial knowledge is gained.

---

[28]   This is not to say that no such view can be given at all, merely that it is likely that similarly structured cases will be applicable to a range of different views of knowledge, and what is said. Thus, a wide range of theorists will be committed to rejecting the knowledge of what is said requirement on testimonial knowledge

Chapter Three: Testimony and the Epistemic Uncertainty of Interpretation.

0. Introduction

In chapter two it was argued that testimonial knowledge does not require knowledge of what is said, so problems with content recovery do not necessarily block testimonial knowledge. However, there are different ways for content recovery to be rendered problematic, and it is still possible that some problems with content recovery often do block testimonial knowledge. My aim in this chapter is to outline one such problem (which I call the 'recovery problem'), and explain why the most obvious solutions to the problem fail. The problem is roughly as follows: In many cases of context sensitivity audiences will not be in a position to know which precise proposition was intended by the speaker, and the speaker will not be in a position to know precisely which proposition the audience will recover. This is due to the contextual factors which determine particular resolutions of context sensitivity being too fine grained (or hidden in other ways) for audiences and speakers to quickly and reliably resolve such context sensitivity correctly. As a result, in many recovery problem cases the beliefs audiences form will fail safety and sensitivity conditions on knowledge. This appears to hold even in cases in which the speaker is a reliable belief former (with respect to

the topic of their testimony) and has no intention to deceive.

I start by outlining the recovery problem, and differentiating it from several related problems in the philosophy of language. Next I explain how the recovery problem relates to (and is a serious problem for) the theory of testimonial knowledge. The relation is not as straightforward as one may think, since the recovery problem concerns audiences beliefs about what testimony has been offered, whereas most theories of testimonial knowledge concern only the testimony which actually has been offered. Finally, in the second half of the chapter I consider and reject a series of responses.

However, before continuing it is worth briefly mentioning a related worry raised by Goldberg (2007). Goldberg considers cases in which audiences either recover a content other than what is said, or are at risk of doing so. Call these cases of mismatch. The problem I raise also concerns cases of mismatch. However, Goldberg thinks that mismatch always blocks testimonial knowledge. He gives the following reasoning:

> 'Suppose S tells H that p, but that for some reason or other the process by which H recovers the proposition attested to is not reliable. In that case, even if H correctly recovers the content of S's telling, this process of recovery will involve a knowledge-undermining element of luck. To see this, suppose that the content of S's actual telling was that p. There are nearby worlds in which what S told H was something else – that r, say. In that case, so long as p is a contingent proposition, in some of these worlds p will be false. Even so, in a good many of those worlds in which p is false, H will accept p, taking this (incorrectly) to be the upshot of what S said.' Goldberg, (2007), 44.

Goldberg is mistaken to claim that so long as *p* is contingent there will be nearby worlds at which it is false. Chapter two clearly illustrates that Goldberg's argument fails, at least when speakers and hearers do happen to coordinate on the same content. However, Goldberg's claim seems false for some cases where speakers and hearers fail to coordinate on the same content as well. For example, *p* and *r* might be truth conditionally equivalent and just differ in Fregean sense. Or, if the speaker knows that she may be misinterpreted as asserting *p*, but knows that *p* is true and and close to *r*, then she might allow the audience to misinterpret her. In such a situation it seems that the audience's

belief is a candidate for knowledge. We can also imagine that in some cases the proposition the speaker asserts will be related to the proposition the audience recovers in such a way that they only diverge in truth value at distant worlds[29]. Indeed, in general the mere contingency of a proposition doesn't entail that there are nearby worlds at which that proposition is false. Moreover, several well motivated theories in the philosophy of language and psycholinguistics (for example, Bezuidenhout (1997), Recanati (2004), Heck (2002), Carston (2002), and Sperber and Wilson (1986)) seem to entail that mismatch is extremely common (this was illustrated in the brief discussion of mismatch arguments in chapter one). If this is the case then we would be forced to either adopt a rather widespread skepticism about testimony, or reject these seemingly well motivated theories. I am hesitant to reject such theories without a detailed investigation into the precise ways in which mismatch is epistemically problematic. This chapter constitutes part of such an investigation. I illustrate a particular way in which mismatch often does seem to block knowledge. However, it is far from clear that all cases of mismatch will block knowledge. Thus, although the problem I raise does sanction a sceptical view of many of our seemingly testimonial beliefs, it is not clear that it sanctions the sort of widespread scepticism necessary to reject otherwise well motivated theories of linguistic communication.

1. Context Sensitivity.

In order to get clear on the recovery problem it will be instructive to consider a slightly broader range of issues. This will be useful both in distinguishing the problem from several related issues, and also in highlighting the fact that it is one amongst a web of related issues in the philosophy of language, all of which may raise difficult questions in the epistemology of testimony.

Normal Communication is rife with context sensitivity. That is, often assertions of the same sentence in different contexts will communicate different propositions. As a result, audiences must rely on their knowledge of the conversational context to understand the speaker. There is disagreement about the form this context sensitivity takes. Many theorists identify speech act content with semantic content, accounting for context sensitivity by maintaining that the context

---

[29] Goldberg considers similar responses, and maintains that if the audience believes anything other than the proposition asserted then their belief is not properly testimonial, since it does not involve the correct sort of reliance on the speaker. I am happy to accept this consequence. As I explain in section 2 the problem I raise concerns beliefs audiences form whilst attempting to form testimonial beliefs. It does not matter for the discussion here whether such beliefs are properly testimonial.

invariant semantic content of a sentence is often insufficient to determine a fully propositional content. For the majority of this chapter I will be assuming this view. However, this assumption is not essential, and it is not universally held. Some theorists maintain that semantic content and speech act content come apart. For example, Borg (2012, p14) maintains that the sentence 'there is nothing to eat' semantically expresses the proposition that there is nothing to eat (in some unrestricted domain), but that that sentence will typically be used to assert that there is nothing to eat in some restricted domain (e.g. in the fridge). This distinction between speech act content and semantic content has been embraced by semantic minimalists (e.g. Borg (2004, 2012), Cappelen and Lepore (2004)), who claim that outside of a small set of cases the invariant semantic content of a sentence does determine a propositional content. By drawing this distinction they are able to maintain that speech act content is highly context sensitive whilst denying that semantic content is context sensitive.

In this chapter the recovery problem is framed in terms of an audience's ability to reliably assign the correct value to a context sensitive term (thus, a close relationship between speech act and semantic content is assumed). However, what really matters is that asserted content is highly context sensitive. When we acquire testimonial beliefs we come to believe the proposition we take the speaker to have asserted[30]. The problems which are raised here concern the recovery of asserted contents. So, for example, when this chapter discusses the difficulties of assigning the correct value to a context sensitive term it may be controversial whether that term really is semantically context sensitive. However, those who deny the semantic context sensitivity will often agree that its use is typically context sensitive. Thus analogous problems will arise concerning the audience's recovery of the speaker's intended meaning. Thus, although the framing of the problem assumes the falsity of semantic minimalism, the main point holds for minimalist views and contextualist views alike, as long as it is acknowledged that asserted content is context sensitive (we will consider minimalism about asserted contents in section 5).

The context sensitivity of asserted contents raises several important questions. There is a question regarding which additional factors determine the proposition expressed by an assertion of that sentence in context. Moreover, there is a related epistemic question concerning how and when

---

[30] Cappelen and Lepore hold that minimal propositions are always asserted along with many other propositions. On their view we might well take the minimal content to be asserted, however we would not treat it as the only, or even the primary content of the assertion.

audiences are in a position to know what has been asserted (and a related psychological question concerning the actual mechanisms by which audiences recover that content).

In the case of simple indexicals (such as 'I', 'here', and 'now') these questions do not seem especially challenging. Although the value assigned to 'I' varies with context, its value is seemingly determined by a simple rule which is, in most cases, easy to apply[31]. However, the metaphysical and epistemic questions become more challenging when we move beyond simple indexicals. Consider demonstratives such as 'that', and pronouns such as 'she'. It is not clear that there is a simple rule (like "'I' refers to the speaker') which relates an occurrence of a pronoun or demonstrative to its referent. In the case of pronouns there are certainly constrains on suitable referents (for example, 'she' can usually only be used to refer to a female). However, these constraints alone are not sufficient to determine a referent. Some further contextual factor must supplement the meaning of the term in order for a referent to be assigned. I follow King (2013, forthcoming a, forthcoming b) in calling terms which require such contextual supplementation 'supplementives'.

Most views of the meta-semantics of context sensitivity raise epistemic questions. However, when the value to be assigned is simply a referent (as in the case of demonstratives and pronouns) these epistemic questions will usually be fairly easy to resolve, for there will not usually be a proliferation of candidate referents to choose from. However, more difficult epistemic and metaphysical questions arise when we consider supplementives with more complex contextual values (call such terms 'complex supplementives'). Consider quantifier domain restriction: usually when someone utters 'every beer is in the fridge' they do not intend to communicate that every beer in the universe is in the fridge. Rather, they intend to communicate that every beer in some more restricted domain is in the fridge. The context must supply a domain restriction, and it is standardly thought that this restriction is supplied in the form of a property (Stanley & Szabó (2000))[32]. In the case of 'every beer is in the fridge' the restriction might be a property such as '..for the party'. So the proposition expressed might be 'Every beer *for the party* is in the fridge'. The problem with such contextual values is that there will usually be many very similar values which could be assigned in a given instance.

---

[31] Although, see Sidelle (1991), Predelli (1998a, 1998b, 2002, 2011), Romdenh-Romluc (2002, 2006), Corazza, Fish, and Gorvett (2002), Gorvett (2005), and Cohen (2013).

[32] Stanley and Szabó explain that only a property (or other intensional restrictor) will give the correct modal profile for the proposition expressed.

This gives rise to both metaphysical and epistemic questions. For example, imagine you arrive at a party and I say 'every beer is in the fridge'. Is the correct restriction 'every beer *for the party* is in the fridge', 'every beer *I am free to offer* is in the fridge', or some other proposition? Whatever one's meta-semantics for supplementives it seems unlikely that one will be able to easily account for the fact that one of these restrictions is correct and the others are incorrect. For example, it seems unlikely that a speaker would intend one of these restrictions over any of the others[33]. This suggests that it may often be indeterminate what is asserted. This indeterminacy about what is asserted raises interesting questions for the epistemology of testimony, since theories of testimonial knowledge are usually formulated in terms of a speaker having asserted (or said, or testified that etc.) a particular proposition. However, I will not be focusing on these questions here, rather I will be focusing on a purely epistemic problem.

Assume for the moment that speakers do generally assert particular propositions, and that audiences aim to recover the proposition asserted[34]. An epistemic problem arises. When complex supplementives are used there will be a multitude of potential values which could be assigned. Some of these values will be extremely similar. Thus, the contexts which determine one proposition over another will differ only very minimally[35]. The features of such contexts which determine that one proposition rather than another is asserted will be extremely fine grained. Indeed, such contexts will often be close to indiscriminable[36]. This will be true regardless of one's meta-semantics of context sensitivity. Unfortunately, the audience's knowledge of the context is comparatively coarse grained. Audiences interpret extremely quickly, and without much conscious consideration of the evidence (especially in low stakes or casual situations). It would take more time and processing power than a normal audience has at their disposal to make the sorts of very fine grained distinctions required to eliminate all the competing interpretations. The situation seems even more severe if one adopts the speaker intentions view assumed in chapter two: speaker

---

[33] The dominant view is that a speaker's intention that a particular value be assigned is at least a necessary condition on that value being assigned. See Donnellan (1966, 1968), Kaplan (1989), Åkerman (2009, 2010), Stokke (2010), and King (2013, forthcoming a, forthcoming b).

[34] This assumption seems common in the epistemology of testimony. However, I will argue at the end of section 5 that we can do away with it and still get the problem.

[35] Its unclear which aspect of the context determines the proposition expressed. The default view is that speaker intentions determine the asserted proposition. Thus, the 'fine grained aspects of the context' which determine different asserted contents might be slight differences in the speaker's intentions. The recovery problem will thus concern our ability to discern contexts at which speakers have slightly differing intentions.

[36] This argument is similar to what Dorr and Hawthorne (2014) call an 'argument from abundance'. A similar line is pushed by Williamson (1997).

intentions are internal and only revealed via the linguistic clues they provide, thus it is not clear that any amount of time or processing power would allow the audience to select the correct value[37].

So, fine grained aspects of the context determine the values complex supplementives are assigned, yet the audience's knowledge of the context is comparatively coarse grained, meaning that there will often be multiple propositions which, for all the audience knows, could have been asserted (call these propositions 'epistemic candidates for what is said' or just 'epistemic candidates'). These epistemic candidates will not be constrained by the speaker's communicative intentions (or any other relevant aspect of the context) since they are precisely candidates which, for all the audience knows, may have been intended (or determined by some other aspect of the context). That is, even if the speaker's intentions suffice to determine that one proposition is asserted, the audience's contextual knowledge may not be sufficient for them to identify precisely which proposition this is (as the audience's access to the speaker's intentions will be limited). Since we are assuming that audiences do reach a somewhat determinate interpretation, the psychological processes underlying understanding must select one of these epistemic candidates over the others. Although these processes will no doubt reliably track epistemic candidates, there is no reason to think that they would reliably result in the audience entertaining the precise proposition asserted, for it is unclear how they could do so.

In order to get clear on the problem it is worth considering an example. Matt and Sally are at Matt's house. Sally says 'I am hungry, is there any food?', to which Matt responds 'Sorry, there isn't any food, lets order a pizza'. Sally agrees, and they order a pizza. Clearly Matt was not saying that there isn't any food *anywhere.* He was saying that there is no food in some restricted domain. However, there are many similar ways the domain could be restricted, here are a few:

    1. There isn't any food *belonging to Matt.*

    2. There isn't any food *belonging to Matt or Tom (Matt's house mate).*

    3. There isn't any food *that Matt is willing to share.*

    4. There isn't any food *which Sally likes and which meets the above criteria.*

---

[37] Indeed, supposing that in order to rationally intend to communicate *p* a speaker must have a reasonable expectation that the audience will recognise that intention, it seems irrational for a speaker to intend any particular value, for the audience will rarely ever be in a position to recognise which particular value was intended. This leads Buchanan (2010) to conclude that propositions cannot be the objects of speaker meaning.

5. etc[38].

How could Sally reliably conclude that one of these restrictions was is correct? In a typical situation Sally would process and respond to the assertion extremely quickly, without explicit reflective consideration of the evidence. She would likely assign a restriction similar to the ones listed above, but she would be lucky if she assigned the precise correct one. This is true regardless of Matt's communicative intentions (or any other objective features of the context) since Sally's evidence (in a typical situation) would not allow her to distinguish between a context in which, say, 1 was intended (or otherwise determined) rather than 2.

This concludes my basic outline of the recovery problem. Before showing that the recovery problem generates a problem in the epistemology of testimony it will be necessary to say a word about epistemic norms in order to get clear how exactly the recovery problem relates to current views of testimonial knowledge and justification.

2. Testimony and Uptake Norms.

The recovery problem primarily concerns our ability to reliably recovery what is said. However, as noted in the introduction the conditions for testimonial justification or knowledge are usually formulated in terms of a speaker having said that *p* and certain other conditions being met. For example, most theorists accept something at least as liberal as the following[39]:

> Uptake: A hearer *H* has the epistemic right to accept speaker *S*'s testimony that *p* if (i) there are no (doxastic or normative[40]) defeaters, and (ii) *H* monitors *S* for trustworthiness (and is thus sensitive to the speaker's trustworthiness).

---

[38] It might be thought that the correct restriction is 'relevant food', or 'available food'. However, it is not clear what actually counts as available or relevant. Does food that Sally dislikes count as relevant or available? It is not clear, thus the epistemic problem will still occur on such restrictions.

[39] With anti-reductionists usually accepting a far more liberal norm which doesn't require condition (ii) (or a similar condition) to be met.

[40] A normative defeater is a proposition an agent *should* believe, which would undermine their belief.

Thus, it may not be clear how the recovery problem relates to the epistemology of testimony. After all, the recovery problem simply illustrates that sometimes audiences will end up recovering a proposition other than what is said. As formulated above, norms such as Uptake do sanction belief in those recovery problem cases where the audience does happen to recover what is said[41], so it is worth considering the epistemic status of the beliefs formed in such circumstances. However, most uptake norms are silent on what the audience should do in the majority of recovery problem cases, since it is unlikely that audiences will recover the precise proposition asserted in such cases.

In order to see why the recovery problem is of interest we must briefly consider the way epistemic norms relate to epistemic practices. Epistemic norms often hold that an audience should form a belief only under certain conditions. However (at least for externalist norms) agents are not always in a position to know that those conditions obtain. Thus, the epistemic practices of a normal agent attempting to follow such norms will sometimes deviate from the practices of an ideal agent attempting to follow such norms. If an agent frequently ends up in situations where it misleadingly appears that the conditions for justified belief (or knowledge) obtain then, whilst trying to follow aforementioned epistemic norms, agents will violate said norms and form unsanctioned beliefs. If we find out that agents do (with some degree of regularity) find themselves in such situations then it is worth considering the epistemic status of the beliefs thereby formed. If such beliefs fail to constitute knowledge then we will have identified a set of beliefs normal, seemingly responsible agents (ourselves included) tend to form in particular circumstance, which fail to constitute knowledge. Thus, we will be warranted in adopting a more sceptical stance to beliefs formed via that method in those circumstances.

This is precisely the situation which arises with respect to the recovery problem. Although the epistemic norms endorsed by most epistemologists don't sanction belief in the majority of recovery problem cases, normal epistemic agents attempting to form beliefs in accordance with such norms will form unsanctioned beliefs in cases where the recovery problem arises. If such beliefs fail to constitute knowledge then we will have discovered a set of unremarkable circumstances in which seemingly responsible agents regularly go wrong when attempting to acquire testimonial knowledge.

---

[41] This result could be avoided by adding a condition requiring that audiences know what has been said. In the objections and responses section I will argue that this response is unsatisfactory.

3. The Recovery Problem and Testimonial Knowledge.

In order to see why the recovery problem often blocks knowledge we must consider not only the audience's knowledge of context, but also the speaker's knowledge of context. An idealised speaker would know which interpretations they leave open to the audience. Thus, a trustworthy idealised speaker would only leave open epistemic candidates which they knew to be true. So, testimony from such a speaker would usually yield safe and sensitive beliefs even in cases where the recovery problem occurred. Unfortunately actual speakers fall short of this ideal. Ordinary speakers are limited in many of the ways that audiences are. The psychological factors that lead an audience member to select one epistemic candidate over another are internal, thus the speaker will not know which interpretation the audience will select. Moreover, like audiences speakers have neither the time nor the cognitive resources with which to consider all the interpretations they may be leaving open. Rather, speakers will often make an assertion on the basis of their knowledge of a particular proposition, with little consideration of ways in which the audience might misinterpret them. The extent to which speakers monitor for potential misinterpretation will, of course, vary with context – when the message is extremely important and misinterpretation carries a high degree of risk speakers will be far more careful about what they say. However, the majority of our everyday communicative interactions are not like this. We assert quickly and move on. This is corroborated by the studies by Keysar et al discussed in chapter one. Those studies indicated that speakers generally over estimated their ability to reliably communicate ambiguous information. If this is correct even in experimental conditions where participants are made explicitly aware of the fact that they are communicating with ambiguous terms then it seems likely that speakers are even less sensitive to small errors in comprehension in normal cases. Thus, in recovery problem cases speakers will often run the risk of leaving open interpretations which are false, or not known to be true.

Consider a case where the speaker does inadvertently leave open some false epistemic candidates. In such cases the audience will select from a group of propositions, some of which are true, and some of which are false. The cognitive mechanisms which lead them to select one interpretation over another will not be sensitive to the truth of the epistemic candidates (it is hard to see how they could be). Thus, if the audience selects a true proposition from amongst the epistemic candidates it

will largely be a result of luck. To illustrate this, consider Matt and Sally again: Matt left open the following epistemic candidates:

    1. There isn't any food *belonging to Matt.*

    2. There isn't any food *belonging to Matt or Tom (Matt's house mate).*

    3. There isn't any food *that Matt is willing to share.*

    4. There isn't any food *which Sally likes and which meets the above criteria.*

    5. etc.

He might have asserted 'There isn't any food' on the basis of his knowledge of any of these. Imagine that he asserted it on the basis of his knowledge of 1. Would the falsity of, say, 2 prevent him from asserting 'There isn't any food'? In many cases it would not. Suppose that Tom does have some food, and that he is generally very open about sharing his food. In such a circumstance 2 would be false. Yet its falsity would not reliably block Matt's assertion because, given the limited time and cognitive resources he has at his disposal whilst planning his utterance (together with the inattentiveness which typifies casual low stakes utterances), it could easily fail to occur to him that a reasonable audience might interpret him this way. Yet, an audience member who knew about Tom's liberal attitude toward sharing may well interpret Matt this way, especially if Tom is salient to the audience at the time of interpretation. In such a case the audience would form a false belief. Indeed, even if they selected a true epistemic candidate it would largely be down to luck, since they could have easily selected the interpretation determined by 2 instead. In such a case belief formation is reminiscent of pulling propositions out of a hat containing both true and false propositions. The psychological mechanisms which guide interpretation will no doubt ensure that the hat contains no obviously false propositions. However, it is hard to see how they could filter out propositions like 2 in cases where they are false.

So, in some recovery problem cases the speaker will leave open false epistemic candidates, and the audience will not gain knowledge. But what about cases in which all the epistemic candidates are true? It appears that these cases are also problematic. Consider a version of the Matt and Sally case in which every epistemic candidate is true; as it happens there is no food at all in the house. Suppose that Sally comes to the true belief that there is no food owned by Matt or Tom, but Matt

made his assertion on the basis of his knowledge that he himself had no food. In such a situation Sally's belief will fail both sensitivity and safety conditions on knowledge. Consider Sensitivity first:

> Sensitivity: 'A belief by S that p is 'sensitive' iff were it not so that p, S would not believe that p'. Sosa, (1999), 141.

That is, a belief that p is sensitive iff in the closest possible worlds where p is false S no longer believes that p. Sally's belief that there is no food belonging to Matt or Tom fails this condition. Many of the closest worlds at which it is false, for example worlds where Tom has some food, are worlds at which Matt still says 'there is no food'. After all, he asserted it only on the basis of his knowledge that he had no food, he never considered Tom's food. Thus many of the closest worlds at which p is false are worlds at which Sally still comes to believe p. Next consider Safety:

> Safety: 'A belief by S that p is 'safe' iff: S would believe that p only if it were so that p'. Sosa, (1999), 142.

Sally's belief also fails safety (at least, in many cases). Supposing that it is normal for Tom to keep food in the house it could easily have been the case that the proposition Sally came to believe was false. That is, supposing that Tom normally keeps food in the house, there are plenty of nearby worlds in which he does have food. However, since Matt asserted only on the basis of his knowledge that he himself lacked food (without ever considering Tom) he would have still uttered 'There isn't any food' in these worlds. Moreover, since the two situations are phenomenologically indistinguishable to Sally she will form the same belief (that neither Matt nor Tom have any food) in these worlds. Thus, there will be a significant number of nearby worlds in which Sally forms the same belief via the same method, in which her belief is false. More generally, in recovery problem cases in which the proposition the audience comes to believe differs slightly from the proposition the speaker intends, there will often be nearby worlds in which the audience forms the same belief but in which the belief is false. That is, in such cases the truth of the proposition believed by the audience will not be strongly tied to their reason for belief - the speaker makes an utterance on the basis of their knowledge of a proposition p_1, the truth of which is independent of the proposition p_2

which the audience comes to believe. It will often be a matter of mere luck when both propositions turn out to be true[42].

It is worth noting that this problem does not depend on any particular view of the meta-semantics of context sensitivity. Essentially the problem is that in certain cases speakers will be disposed to utter a sentence S on the basis of their knowledge of a proposition p (and with the intention to communicate p), in circumstances where a similar proposition q is false, and in which some competent audiences might interpret an utterance of S as expressing q. If one has an intention based meta-semantics then this will be a situation in which a speaker asserts p and is misinterpreted. However, if one thinks that what is said is audience sensitive then it might be a situation in which the speaker accidentally asserts q, or in which it is indeterminate what is actually said. As long as the audience thinks something has been asserted, and the speaker takes themselves to have made an assertion, the problem can occur.

It is an empirical question just how often the recovery problem arises. It will likely arise in many contexts where complex supplementives or loose talk are present, and in which speakers are not optimally attentive to possible misinterpretation. To get a grasp on just how common such situations seem to be we need merely reflect on the sort of situations we find ourselves in every day when we socialise at the pub, relax at home with our partners, or engage in passing small talk in the department. These situations make up a significant percentage of our communicative interactions, yet they are precisely the sorts of situations in which we speak loosely and reflect little on how we might be precisely interpreted. Thus, the recovery problem is not confined to a small class of cases we can safely ignore[43].

---

[42] The situation gets worse when we consider chains of testimony, which have the potential for a Chinese whisper effect. Even if the first audience member entertains the proposition the speaker intended, it is unlikely the final member of the chain will. Whatever epistemic value there was in the first testimonial belief will be diluted further down the chain. Things are worse still if a speaker is required to know p in order for an audience to gain knowledge that p (although, see Lackey (1999)), for a single problematic utterance could cause all the following testimonial beliefs to fall short of knowledge.

[43] It seems that the problem is less likely to arise in high stakes contexts (where misinterpretation carries a risk of meaningful repercussions). There are two reasons for this. Firstly, hearers accept testimony less readily in such contexts (and hedge their beliefs more). Secondly, speakers are more careful about what they say in such contexts.

4. Response: Defeaters.

Uptake contains a no (normative or doxastic) defeater condition. Thus, if there are doxastic or normative defeaters present in recovery problem cases then audiences attempting to behave in accordance with Uptake will not form beliefs. This response faces several problems.

Firstly, normative defeaters are facts or propositions which defeat an agent's belief and which the agent *should* be aware of. This suggests that agents are blameworthy for not being aware of normative defeaters. If this is correct it seems that if a fact is to count as a normative defeater then it should be something an ordinary agent can be reasonably expected to grasp. If we have to do a lot of theoretical work in order to discover *p*, work which the average agent is not in a position to do, then *p* is not a normative defeater. However, the recovery problem, and the problems it causes for testimonial belief formation, have not been widely recognized by even epistemologists of testimony. If the experts on testimony have not recognized the phenomenon or its problematic nature then it seems far fetched to claim that average audience should. Thus it seems far fetched to argue that there are normative defeaters available in recovery problem cases.

Secondly, it is not clear how normal audiences would come to possess doxastic defeaters which would block belief in recovery problem cases. The most plausible approach would be to follow Lackey (2006) in claiming that audiences possess inductive grounds for considering certain conversational contexts to be epistemically unsafe, and argue that these contexts include the sorts of situation in which the recovery problem is likely to arise. Lackey is surely correct that audiences posses inductive grounds for considering certain contexts to be unsafe. For example, we are far more trusting of assertions made in the doctors office than at the poker table. Moreover, it does seem that the recovery problem will be more prevalent in certain types of context. These will be contexts in which the stakes are low, and the information communicated is relatively unimportant. In such contexts loose talk and context sensitivity will be more common, and speakers will dedicate fewer resources to checking for potential misinterpretation. Do audiences possess evidence that such contexts are epistemically unsafe? It is hard to see how they would. We rarely become aware of minor miscommunications. In the sorts of situations I have been discussing the conversation will usually carry on smoothly despite the slight miscommunication. Indeed, it might even be thought that audiences possess positive (misleading) reasons for thinking such contexts are unproblematic.

The propositions left open in most recovery problem cases will usually be close to the truth, especially in relevant practical consequences[44]. Thus, when agents act on their testimonial beliefs it will be, in most recognisable respects, as if those beliefs were true. This generates the illusion that low stakes contexts are epistemically safe environments (a similar point is made by Keysar (2007) who points out that, from all we know by engaging in normal communicative behaviour, linguistic communication could actually be significantly less reliable than we generally assume).

Indeed, as discussed in more detail in chapter one, and as mentioned earlier in this chapter, there is even experimental evidence that we overestimate our own effectiveness as communicators. Boaz Keysar and Anne Henley (2002) found that speakers regularly overestimate their own communicative abilities when making ambiguous utterances. Even when made aware of the specific ambiguity in their utterances speakers were found to significantly overestimate the reliability with which audiences were able to recover their intended message. If such overestimation of reliability is significant even under experimental conditions where the participants are made explicitly aware of the possibility of miscommunication, then it is likely even more prevalent in ordinary contexts where the possibility of miscommunication is not made salient. Thus, the defeater strategy does not seem promising.

A related response is that uptake norms should be formulated with the requirement that the audience be in a position to know that p has been asserted (or intended). In chapter two we saw that the knowledge of what is said requirement on testimonial knowledge should be rejected. However, there are two further problems with this as a response specifically to the recovery problem. Firstly, it usually appears to us that we know what the speaker has said (even in cases of loose talk and context sensitivity). Thus, audiences attempting to follow such modified norms would still form beliefs in recovery problem cases. Secondly, the sceptical result emerging from this view is actually even more extreme than the sceptical view which seems to arise from the recovery problem as I have formulated it. I argue that audiences often lack knowledge in recovery problem cases. However, if we were to accept the 'knowledge of what is said' requirement then we would be forced to claim that audiences are never even justified in forming testimonial beliefs in any recovery problem cases – even when the speaker knows which epistemic candidates have been left open, and knows each candidate to be true.

---

[44] What I mean by this is that given the aims and interests of the conversational participants similar courses of action will usually be sanctioned by each epistemic candidate. True and false candidates alike. This is the sense in which they are close to the truth.

5. Response: Alternative Propositional Objects of Assertion[45.]

So far it has been assumed that speakers generally intend to assert, and audiences generally recover, single fine grained propositions. This assumption can be questioned. Perhaps the testimonial beliefs we form, and are justified in forming, are coarse grained. If the propositions audiences recover are highly coarse grained then we might be able to avoid the recovery problem. This will be the case if the propositions believed are true at all (or almost all) the worlds at which any of the fine grained epistemic candidates are true (such a proposition would be equivalent to the disjunction of the fine grained epistemic candidates). Beliefs in such propositions will not be as vulnerable to failures of safety and sensitivity. This response gains intuitive support from the fact that we don't seem to explicitly form the sorts of fine grained beliefs I have been discussing. For example, thinking back to the Max and Sally case, it seems unrealistic to claim that Sally explicitly entertains the restriction "food *which belongs to either Matt or Tom*".

There are three ways of developing this response: Firstly we might maintain that audiences only form a very general testimonial beliefs. Secondly, we might maintain that the audience comes to believe a coarse grained proposition in addition to some more fine grained proposition. Finally, these responses can be combined by maintaining that we only form testimonial beliefs in very general propositions, but merely accept, or assign raised credences to, fine grained propositions.

Before discussing these responses it is important to consider the intuitive claim that we don't form the sorts of fine grained beliefs that give rise to the recovery problem. I think that it is true that we don't *explicitly* form such beliefs. However, this does not undermine the problem. To see this, consider the commonly drawn distinction between occurrent (or explicit) and dispositional beliefs (a similar distinction is drawn by Buckwalter, Rose, and Turri (forthcoming)). Often when we form a belief we explicitly represent the proposition we come to believe, and mentally assent to it. Clearly Sally would not explicitly represent and assent to the proposition 'there isn't any food *belonging to Matt or Tom*' in response to Matt's assertion. Indeed, her explicit mental representation may be somewhat indeterminate in the same way Matt's assertion is. However, the mere fact that

---

she does not explicitly represent and assent to the proposition does not mean that she does not believe it. Many of our beliefs are merely dispositional, or implicit. For example, you no doubt believe that there are no plastic mice on Mars. However, it is unlikely that you ever explicitly represented and assented to this proposition (until now). Likewise, if I see what appears to be a racing car approaching I might explicitly represent and mentally assent to the proposition 'there is a racing car approaching'. Yet, it is unlikely that this exhausts the beliefs I form upon seeing the car. For example, although I may not explicitly represent and assent to the proposition 'that car will be loud', I might still cover my ears, or exhibit surprise if it were to pass by without making a sound. The mental state that gives rise to these dispositions (or perhaps consists in these dispositions) is formed in response to the evidence of my senses, and it interacts with my desires (for example my desire not to be deafened) in order to bring about action. So, although I never explicitly entertain the thought 'the car approaching will be loud', it certainly seems that I have a belief along those lines. My claim is that we often form relatively fine grained dispositional beliefs in response to testimony. These beliefs are revealed through the dispositions we acquire when confronted by another's testimony.

It is conceivable that one might wish to save the word 'belief' for mental states involving explicit representation of, and assent to, propositions. We might label unconscious mental representations belief*. Such a change would not make a substantial difference. My belief* will represent the world as being a certain way, this representation will guide my behaviour, and it will be evidence responsive[46]. Thus, it is hard to think of a reason to worry about unjustified, false, or unsafe beliefs which are not also reasons to worry about unjustified, false, or unsafe beliefs*. If the fine grained representations I am claiming we acquire in response to testimony are in fact beliefs*, then the recovery problem will be just as worrying. For that reason I will continue to call these unconscious mental states simply 'beliefs'.

The first version of the coarse grained proposition response maintains that the only beliefs normal audiences form in recovery problem cases are highly coarse grained. This, entails that it would be unusual for Sally to form a belief such as 'there is no food *belonging to Matt or Tom*'. Of course, it would be unusual for her to explicitly form such a belief. However, we have just seen that not all the beliefs we form are fully explicit. Many of our beliefs are reviled through the dispositions we

---

[46] Indeed, as Rose and Schaffer (2013), and Buckwalter, Rose, and Turri forthcoming argue, they are capable of constituting knowledge.

acquire upon forming them. If Tom's open food policy were salient to Sally at the time of Matt's assertion then she would, in many cases, acquire the disposition to be surprised upon finding her favourite foodstuff in the fridge labelled 'Property of Tom'. This suggests that it would not be unusual for her to form a dispositional belief concerning Tom's food, meaning that standard practice does not involve *only* forming highly coarse grained beliefs. Such a reaction would be unusual if we normally form only highly coarse grained beliefs in recovery problem cases. So this response fails[47].

A second way to push this objection is to maintain that audiences form highly general beliefs in addition to more fine grained beliefs. Since these coarse grained beliefs will usually be safe and sensitive audiences will usually gain some knowledge in addition to some unsafe fine grained beliefs. It is unclear whether this constitutes really a response at all. We retain the consequence that ordinary agents form many false or luckily true beliefs in recovery problem cases, and merely add that we do at least gain some minimal knowledge in addition. It is not clear that the additional knowledge adds much of value, since we will still act on our more fine grained beliefs.

To get a feel for just how unsatisfying this response is consider an analogous skeptical argument concerning perception. Suppose it is argued that many of the perceptual beliefs we form in certain circumstances are the result of potentially misleading cognitive penetration (the influence of expectations, beliefs, and biases on the content/character of one's perceptual experience), and are at best luckily true (see Siegel (2012, 2013, and forthcoming) for discussion of the actual epistemic impact of cognitive penetration). One might argue that my current belief that there is a water bottle in front of me is likely to be either false or luckily true because, as a result of cognitive penetration, I would have had the belief that there was a water bottle in front of me even if it were in fact a beer bottle. In such circumstances I would still form the true beliefs that there is a bottle in front of me, and that I would be refreshed if I were to consume its contents. However, this observation should provide little comfort, since I still have a fine grained false or luckily true belief which I am likely to act on. Likewise, it is of little comfort to maintain that Sally at least gains *some* testimonial knowledge, since she also forms a significant belief which she may well act on, and which falls sort of knowledge.

A third version of the coarse grained proposition response runs as follows: the object of testimonial

---

[47] Additionally, there will still be a proliferation of very similar coarse grained epistemic candidates. Coarse grained propositions can differ in fine grained ways. It is not clear how the audience could reliably recover the correct coarse grained proposition over one of the many similar coarse grained propositions, thus the recovery problem will still arise.

belief is coarse grained, so testimonial beliefs are safe and sensitive. However, we adopt weaker attitudes to the fine grained epistemic candidates. These weakened attitudes explain the dispositions we acquire as a result of uptake. There are two ways to develop this line – either in terms of mere acceptance (perhaps as discussed in Bratman (1992), or Van Fraassen (1980)), or in terms of a raised (but sub-belief) credence. The main reason to be skeptical of the mere acceptance response is that acceptance seems to be a reflective attitude – a way of hedging our beliefs[48]. However, we seem far less prone to hedging in sorts of context in which the recovery problem arises (low stakes casual contexts where miscommunication carries little practical risk). Uptake in low stakes contexts is usually fast, automatic, and unreflective. Thus it is unlikely that our default reaction in such contexts is mere acceptance.

The second version of this response does not seem to capture the full extent of the dispositions we acquire on the basis of testimonial uptake. Perhaps a raised (but sub belief) credence is enough to explain Sally's surprise at finding her favourite food in the fridge. However, there are more problematic dispositions Sally may acquire acquire. Consider the following continuation of the Matt ad Sally case: Matt tells Sally 'there isn't any food'. Tom's partner then enters and asks if there is any food. Sally responds 'no, there isn't any food, but we're ordering a Pizza'. This is a fairly natural exchange. However, Tom's food would be considered available to his partner if he had any. This suggests that the mental state which triggers Sally's assertion disposes her to assert as if Tom doesn't have any food. It is unusual to assert on the basis of mere high credences, such actions generally require (at least) all out belief[49]. If Sally's assertion is not unusual then all out belief in the fine grained proposition is not unusual either. Thus the problem remains.

One might worry that this exchange only seems natural because Sally interpreted Tom's partner as asking about food which Sally or Matt were in a position to share, or because Sally might not think about Tom's food when replying. The second interpretation seems unpromising because we have already postulated that Tom, and his food sharing policies, are salient to Sally. With respect to the first interpretation; we can imagine Tom's partner asking 'Does anyone know if there is any food?'. This question is more naturally read as concerning food available to Tom's partner, rather than food Matt and Sally are in a position to share. Yet Sally's response would still not be unusual.

---

[48] This is not merely an intuitive point – there appears to be psychological evidence which points toward our default attitude being belief. For example, Gilbert (1991) argues that when we adopt an attitude like disbelief, or some weaker intermediate state, we start by forming a belief and then weakening it.

[49] Indeed, it has been argued by Mark Kaplan (1996) that to have a belief in *p* simply is to be disposed to assert that p if forced to assert either *p* or not-*p*.

A slightly different version of the coarse grained proposition response runs as follows: the object of testimonial belief is a coarse grained proposition, thus our testimonial beliefs are safe and sensitive. However, we often move from our coarse grained testimonial belief to more fine grained beliefs which fail safety and sensitivity. These beliefs are reached by combining our coarse grained testimonial belief with our background assumptions (and other psychological factors), thus they are not truly testimonial. This response is unsatisfactory, for it seems to be a merely taxonomical point. Whether these beliefs are in the strictest sense testimonial is not very important once we recognise that they are still problematic, and still commonly arise as a result of testimony. Indeed, my claim has been that in many recovery problem cases the proposition recovered is not the asserted proposition any way. So it was antecedently unclear whether or not we should, strictly speaking, call them 'testimonial beliefs' any way.

It is worth considering whether semantic minimalism offers a response to our problem. If the context invariant semantic content of a sentence is fully propositional, then this proposition seems a natural candidate for the object of testimony. For Cappelen and Lepore the minimal semantic content of a sentence is the proposition expressed by every assertion of that sentence[50]. There are two related reasons to be skeptical of the claim that such minimal propositions are the objects of testimony. Firstly, these minimal propositions do not convey much information, and the information they do convey seems to be of limited value. Imagine that Matt and Sally are about to go for a walk. Matt is waiting, and Sally shouts 'I'm ready'. Matt is interested in the information that Sally is ready to go for a walk. However, this is not the minimal semantic content of the utterance. Rather, the minimal content is simply that Sally is ready. This is the same minimal content which would be expressed by an utterance of 'I'm ready' in a context where Sally is ready to convert to Buddhism, or to give up drinking. It is not clear why Matt would be interested in this information, or how his knowledge of this minimal proposition alone would guide his action. Thus, if the objects of testimonial knowledge are minimal propositions, then testimonial knowledge seems ill suited to play the sort of central role we assign it in our every day lives.

Secondly, assuming that testimony is our central means of sharing knowledge it would be hard to explain the importance of labels such as 'liar' if they do not track testimonial wrongdoing. However, as Jennifer Saul (2012) argues, our assessments concerning whether or not someone has

---

[50] Borg (2012) holds that the minimal content of an otherwise sincere truthful assertion is often false, making it a poor candidate for the object of testimony.

lied do not track minimal propositions. Saul asks us to consider Bill Clinton's assertion of 'there is no improper relationship', shortly after his affair with Monica Lewinsky. Technically Clinton was not lying since the affair was over. However, if Joe Blogs, a current adulterer, were to respond the same way to similar allegations he would be lying. Precisely the same minimal proposition is expressed in each case. Thus if our assessments track minimal contents then we should reach the same verdict about each case. Since verdicts about each case differ it appears our assessments don't track minimal contents. If our most central normative assessments of assertions don't track the contents of testimony, but rather some other level of content, then this calls into question the relative importance of testimony to ordinary agents. So, once again, the minimalist response calls into question the central role testimony seems to play in our lives.

I think that all versions of the coarse grained propositions response fail. However, there is a related response which already has some support in the literature on context sensitivity - the view that typical utterances have multiple contents (this view is known as 'speech act pluralism' or 'propositional profusion' and was discussed briefly at the end of chapter two). This view can take several forms, from the view that only extremely similar propositions are expressed (discussed in Dorr and Hawthorne (2014)) to the view that a huge number of different propositions are expressed, propositions which the speaker needn't even be aware of expressing. Such a view is endorsed by Cappelen and Lepore (2004)(2006). Cappelen and Lepore motivate their view by drawing our attention to the different ways in which we report speech, for example, they ask us to consider Richard Nixon's 'smoking gun' utterance:

> 'When you get in these people, when you get these people in, say: 'Look, the problem is that this will open the whole, the whole Bay of Pigs thing, and the president just feels that,' ah, without going into the details…don't, don't lie to them to the extent to say there is no involvement, but just say this is sort of a comedy of errors, bizarre, without getting into it, 'the president believes that it is going to open the whole Bay of Pigs thing up again, and ah because these people are plugging for, for keeps and that they should call the FBI in and say that we wish for the country, don't go any further into this case'. Period. That's the way to put it, do it straight.'

They then note that this single utterance has been reported the following ways:

> 'Nixon told Haldeman to tell the CIA to tell the FBI not to pursue their investigation into the

Watergate Burglary.

Nixon is clearly heard telling his chief of staff, Bob Haldeman, to implement John Dean's idea that the CIA be used to pressure the FBI to limit the Watergate investigation.

Nixon wanted the CIA Director Richard Helms to thwart the FBI's probe of theWatergate Burglary by saying it was a CIA operation.

Nixon told Haldeman to tell Helms that Nixon wanted him to stop the Watergate Investigation.

Nixon told Haldeman to break the law.' Cappelen and Lepore (2006), 1049-1050.

These reports all have different contents, yet each report seems reasonable. So the most straightforward thing to conclude is that all of these contents were asserted. Indeed, speech act pluralism is also one constituent of Cappelen and Lepore's solution to the problem of content sharing, which is closely related to the recovery problem. The problem as they put it is that contextualist views make communication miraculous - we would be incredibly lucky to regularly coordinate on precise contents when communicating with context sensitive terms given the many ways in which we must rely on our knowledge of the context (this argument is discussed further in chapter four). They reject contextualism for all but a very limited set of terms, and argue that there is always a fully propositional context invariant semantic meaning which is shared amongst utterances and which is thus easily recoverable. However, as we have just seen, it is unlikeley that this minimal content exhausts the asserted content. We usually communicate far more than this minimal proposition. This is where speech act pluralism comes in. Many other propositions are also expressed, and the audience's representation of the context determines the propositions they recover. It might be thought that speech act pluralism would offer a response to the recovery problem. For a start one might adopt Cappelen and Lepore's whole view and maintain that in the cases discussed above (for example the cases of quantifier domain restriction) there was a minimal semantic content which is shared, and many other propositions expressed as well. Or one could reject semantic minimalism but still maintain that all the epistemic candidates were in fact asserted. If one took the first line it would be similar to adopting the view that audiences typically form a highly coarse grained belief, but also form a fine grained belief as well. We have already encountered reasons to reject this view. However, if one adopted the latter view one might maintain that when forming testimonial beliefs we do not simply come to believe one single proposition, there are several epistemic candidates, and may even be several propositions expressed, so we come to believe many fine grained propositions on the basis of a single utterance. Some of these beliefs will be problematic (unsafe or even false), but some will constitute knowledge, and that is good

enough. This response faces many of the problems I have already discussed, for example it is unclear the extent to which it is a genuine response - it still has us forming many false beliefs. However, it also has problems of its own. For example, it seems unclear whether *any* of these fine grained beliefs, considered alone, would satisfy the safety or sensitivity conditions on knowledge. Moreover, this response fails if we endorse an agglomeration principle for belief. That is, we run into problems if we maintain that if I form a belief in *p* and a belief in *q* on the basis of a single stimulus then I have formed a belief in (*p* & *q*). If such an agglomeration principle were correct then audiences would typically form beliefs in a conjunction of epistemic candidates, and thus would be at even greater epistemic risk. However, if we deny the agglomeration principle then this response does have one redeeming feature: it establishes that, even in the sorts of context where the recovery problem occurs, testimony produces a baseline of true fine grained beliefs. This baseline of truth may be sufficient to explain why the norms of testimonial belief formation in low stakes contexts have not adapted in response to the epistemic problems which seem to arise in such contexts. More generally however, it is not clear that speech act pluralism can offer a response simply because its only practical consequence for the recovery problem is whether or not we say that the epistemic candidates for what is said are in fact said. Speech act pluralism does not guarantee that only true propositions will be epistemically available to the audience. It merely makes a classificatory point which relates our reporting practices and our ability to share (in some very minimal sense) contents.

So far I have considered several attempts to resolve the recovery problem by denying that audiences come to believe single fine grained propositions (whilst still maintaining that the objects of belief are propositional). However, to be clear, it *is* a simplification to maintain that audiences entertain single precise fine grained propositions. Our communicative intentions, and the testimonial beliefs we form, are undoubtedly vague and imprecise. However, it seems they are not imprecise enough to reliably avoid the problem. This is illustrated by the naturalness with which audiences appear to gain dispositions we would normally associate with more fine grained beliefs. This suggests that some fine grained aspects of the audience's belief state will be relatively settled (settled enough to have an impact on the audience's dispositions), whilst others may be more indeterminate. For example, Sally's belief state may be determinate with respect to the question of whether Matt or Tom have any easily preparable food (for example, frozen pizzas or ready meals), but indeterminate with respect to the question of whether or not they have any food with a medium preparation time (for example, it may be unclear whether a pre-spatchcocked chicken would be considered

contextually relevant). Thus, the recovery problem does not rely on the simplifying assumption that audiences typically form precise/fine grained testimonial beliefs, merely the assumption that the audience's belief states are usually settled with respect to some fine grained questions. And this claim is motivated by the fact that it is far from unusual for audiences to acquire, through uptake, the sorts of dispositions associated with fine grained beliefs.

6. Response: Non-Propositional Objects of Assertion.

Attempts to avoid the problem by postulating alternative propositional objects of assertion (or belief) seem to fail. However, Ray Buchanan (2010) denies that the objects of speaker meaning are propositional, and he does so on the basis of considerations similar to those raised here. So it is worth considering whether his view offers a solution to the recovery problem.

Buchanan claims that the objects of speaker meaning are not propositions, but rather properties of propositions (incomplete propositional templates with vague restrictions as to how they are to be completed). Audiences grasp these properties by entertaining one or more of the propositions which fit the template. Buchanan's view is developed partly in response to an epistemic problem which bears many similarities to the recovery problem: On speaker intention based of meta-semantic views the asserted proposition must be intended by the speaker. In order to rationally intend to assert that p a speaker must have a reasonable expectation that the audience will recognise their intention to assert p. However, due to the proliferation of highly similar propositions which the audience might attribute in cases of context sensitivity there is no single proposition such that the speaker can reasonably expect the audience to recover that precise proposition. Thus, unless speakers are highly irrational, the object of speaker meaning cannot be propositional. However, it would not be irrational for a speaker to expect the audience to recover one (or more) of the propositions in the extension of some vaguely specified property. This is because there is a far greater chance of the audience recovering one of the many propositions which fall in the extension of the intended property.

This problem sounds very similar to the recovery problem, thus one might think that the solution should carry over unproblematically. This would be a mistake. Buchanan's problem concerns the determination of what is said – what is said cannot be determined by speaker intentions if

propositions are the object of speaker meaning. To claim otherwise would be to attribute an unrealistic degree of irrationality to speakers. The recovery problem, on the other hand, does not concern the metaphysics of what is said. The problem is simply that due to the collective epistemic limitations of the speaker and audience speakers will, in certain cases, be insensitive to the ways in which they may be misinterpreted. This is entirely consistent with the notion that what is said is actually a property rather than a proposition. And it is consistent with the view that speakers are not irrational enough to expect audiences to recover a particular fine grained proposition. On a view like Buchanan's recovery problem cases would simply be cases in which either A), the audience is at risk of recovering a proposition which does not fall under the extension of the speaker's intended property, or B) the speaker does not realise that there are propositions in the extension of their intended property which they do not know to be true. It seems likely that the former situation would occur more frequently, unless the speaker were being particularly careless. In such cases the set of propositions in the extension of the speaker's intended property will usually be a subset of the epistemic candidates.

7. Conclusion.

The recovery problem holds that we are often at risk of forming false or unsafe testimonial beliefs as a result of our heavy reliance upon limited contextual knowledge when planning and interpreting assertions. This problem will arise even for agents attempting to follow the epistemic norms endorsed by most epistemologists of testimony. The problem will be more pronounced in low stakes contexts where speakers are more careless and audiences hedge their beliefs less. Several responses have been considered, none of which were satisfactory.

Chapter Four: Testimony, Pragmatics, and Plausible Deniability.

0. Introduction.

In previous chapters it has been argued that communication with context sensitive terms can lead to epistemically problematic forms of miscommunication, which can block testimonial knowledge. In this chapter I explore another problem which seems to arise from these same processes. This problem pertains to the commitments speakers undertake when they make assertions. Because of the many ways the recovery of content can go wrong, speakers are able to make assertions and yet deny responsibility for the proposition asserted, claiming that the audience made a mistake in resolving the context sensitivity. That is, speakers are able to maintain plausible deniability about what is said. Call this the 'deniability problem'. The aim of this chapter is to explain why the

deniability problem is problematic, and start to identify the range of utterances to which it applies.

Elizabeth Fricker (2012) has pushed a similar line of reasoning, arguing that implicatures fail to carry the epistemic force of outright assertions. She argues that speakers maintain plausible deniability about the implicature, meaning that they fail to undertake the commitments necessary to transmit testimonial knowledge. I start by outlining Fricker's view, and arguing that it potentially applies to a far wider range of utterances than those she considers, including many assertions. I then go into greater detail explaining exactly why this is worrying.

The deniability problem is problematic for three reasons. Firstly, many views of testimony (for example, 'telling' based views of testimony such as Fricker (2006a), Hinchman (2005), Moran (2005a, 2005b), and Ross (1986)) emphasise the role of speaker commitments in the justification of testimonial beliefs. The deniability problem entails that speakers often fail to undertake the types of commitments emphasised by such views. On telling based views of testimony, when an audience is told *p* they gain a reason to believe *p* in virtue of the speaker having publicly taken responsibility for the audience's belief that *p*. However, when the speaker maintains plausible deniability about what is said no such commitment is undertaken. Secondly, plausible deniability prevents epistemic buck passing. Sanford Goldberg (2006), and Benjamin McMyler (2013) have argued that the ability to pass the epistemic buck in response to challenges to one's belief is a distinctive epistemic right agents gain only in virtue of forming testimonial beliefs. This diminishes the belief holder's responsibility for their testimonial belief as compared to beliefs formed via other methods. However, when the speaker maintains plausible deniability the audience loses the ability to pass the epistemic buck, and thus fails to gain the epistemic rights distinctive of testimonial knowledge. Finally, the ability to maintain plausible deniability blocks one of the primary disincentives to deceptive or careless assertion.

After explaining why the deniability problem is problematic I focus on identifying the range of utterances to which it applies. I outline a puzzle arising from the recent debate over context sensitivity in the philosophy of language, which seems to suggest that the deniability problem extends to a very large number of utterances. The puzzle is as follows - on the one hand it has been argued that there is widespread context sensitivity in natural language, and audiences must rely heavily on their knowledge of the context to recover the speaker's intended meaning. This includes cases which we would intuitively treat as being on a par with normal testimony. On the other hand, Herman Cappelen and Ernie Lepore (2004) have argued that such widespread context sensitivity

would make communication miraculous, pointing to the same 'problematic epistemics' which, according to Fricker (2012), give rise to plausible deniability. If context sensitivity is very widespread, but has the same 'problematic epistemics' that Fricker identifies for implicature, then the deniability problem is extremely far reaching. This would be a worrying and radical result. Thankfully this result can be avoided, or at least weakened. There have been several contextualist responses to Cappelen and Lepore which aim to establish that widespread context sensitivity does not make communication miraculous. These responses don't work as general solutions to the deniability problem, however they do allow us to limit its scope. We end up with a set of criteria for identifying discourses which are particularly susceptible to the deniability problem. In the conclusion I suggest some important discourses which may still face the deniability problem.

1. The Deniability Problem.

Because of what Fricker calls the 'ambiguous epistemics' of implicature speakers are able to maintain plausible deniability about what they have implied. That is, they maintain the ability to deny that they ever intended to imply what the audience takes them to have implied. Such a denial will involve the construction of a believable narrative in which the speaker's attitudes and expectations at the time of utterance were such that they could reasonably make their utterance without intending to imply what they were taken to have implied. More precisely:

> Plausible deniability - An agent has plausible deniability about intending to communicate a proposition p with an utterance u of a sentence s if that agent is able to tell a story (with at least some degree of believability[51]) about their attitudes and expectations at the time of utterance such that a reasonable agent with those attitudes and expectations could utter s with no intention to communicate p.

The plausibility of the denial will depend on the plausibility of the story about the agent's attitudes. In order to be plausible the story will have to be consistent with what the audience knows about the speaker's attitudes and beliefs. That we have such plausible deniability about implicatures is illustrated by the following example:

---

[51] Plausible deniability is clearly a gradable notion since some denials will be more plausible than others. In section four we will consider some factors which might make some denials more plausible than others.

Implicature: Matt is running out of fuel and needs some fast. He stops and asks a stranger where he can get some fuel. The stranger says 'there is a gas station around the corner'. The stranger thereby implies (implicitly communicates) that the gas station is open and has fuel. However, if Matt were to get to the gas station and find it closed or out of gas he would have a hard time criticising the stranger for her utterance. She could easily maintain that she never intended to communicate that the gas station was open or that it had gas. She can maintain that she was merely suggesting it as a place to try (she could maintain this even if she in fact knew that the gas station was shut or had no gas).

When Fricker talks of 'ambiguous epistemics' she refers to the way audiences and speakers must rely on what she calls 'knowledge context' (or 'K-context') in order to recover the communicated message. Knowledge context is the audience's representation of factors such as mutual knowledge, past utterances, Gricean norms, the mutual goals of the conversation, and any other information generally relevant to interpretation other than basic knowledge of the syntax and invariant semantic content of the utterance. Because of the many complex ways in which audiences must rely on context in order to recover an implicature there are many ways the process can go wrong. For example, the audience could employ an aspect of knowledge context which the speaker never intended them to employ, or they could fail to employ an aspect of knowledge context that the speaker did intend them to employ. Likewise, they could hold false beliefs about the context, or at least beliefs which were not mutual knowledge, and appeal to such beliefs in their recovery of the implicature. Because the recovery of implicatures can go wrong in so many ways the speaker is able to deny responsibility for the proposition the audience recovers, claiming that there was a mismatch between what the audience recovered and what the speaker intended (perhaps even claiming the speaker didn't intend any implicature at all)[52]. Fricker's own take on the problem is as follows:

'Entirely genuine misunderstandings and mistakings are endemically liable to happen, regarding a supposed message that is conversationally implied, not stated, due to the very complex mutual epistemics of the situation. Given these complex epistemics, it is not

---

[52]It has been pointed out by an anonymous reviewer that this problem actually cuts both ways. In some cases contextual indeterminacy seems to grant audiences the freedom to recover the meaning they find most convenient. The reviewer provides the following example : One's partner may say that they are going to have coffee at the local café, and one might interpret them as meaning that they are going to the café *now*, using it as an excuse to drink the last of the coffee. One might do this even if one knows that they might be planning to go later (in which case they may still want some of the remaining coffee). I think this raises some interesting questions about the way in which speech acts make various courses of action permissible. However, there is not enough space to give this issue a proper treatment here.

epistemically feasible to pin undeniable specific commitment onto a speaker: she can always wriggle out of it. This may be in bad faith; but very often it may not—maybe she miscalculated what her audience would infer; maybe she had not really figured it out. Fricker.' (2012), 87.

'Since the claim that she intended to communicate *E* turns on claims about her private intentions and K-context, including her second order representations of others' beliefs, she can always get away with denying that she intended any such thing; even if her denial is made in bad faith. Lies about my own intentions and other mental states may be suspected, but cannot be refuted. In contrast, when someone makes an explicit statement of a fact *P*, what she signs up to in doing so—taking responsibility for the truth of *P*—is a public fact about the situation, determined by semantics and objective features of context. So it cannot be incorrigibly denied by the speaker. I can be nailed as having stated that *P*; never as having insinuated that *P*.' (ibid: 88-89).

Fricker seems to suggest that the deniability problem only applies to implicatures, suggesting that plausible deniability is never available regarding the primary content of one's utterance. The thought is that one can deny that one intended to imply anything without descending into absurdity. However, if one makes an intelligible declarative utterance one cannot, without descending into absurdity, claim that one never intended to communicate anything. This may be true, but one can still maintain plausible deniability about intending the particular proposition the audience recovered without claiming that one never intended to communicate anything at all. To see that speakers do attempt such conversational manoeuvres with asserted contents as well as implicatures consider the following two examples (the first of which occupies a grey area between what is implied and what is asserted, the second of which clearly concerns asserted content)[53]:

Scalar implicature: We are planning a group trip to a theme park and deciding how many cars to take. I wish to cause logistical problems because I hate fun, so I say 'Matt has three

---

[53] As we move through the examples from the clear case of implicature to the case of mere context sensitivity the stories the speaker tells start to sound slightly less plausible. They would certainly raise our suspicions, and if a speaker frequently made manoeuvres like these then we would consider them untrustworthy. This is a point to which I will return when discussing responses to the problem. For now it suffices to note that on one-off occasions we would usually let such matters slide and speakers would usually get away with making such conversational manoeuvres. This is not to say that they wouldn't be criticised for being unclear, simply that they would not be held to what they communicated.

kids' knowing that he has five. On the day of the trip Matt arrives with his five children and we don't have enough space in the car. You challenge me for saying that Matt had only three children. However, I might attempt to maintain that I didn't mean he had *only* three children, I meant that he had *at least* three children, so we would need at least three additional seats. The plausibility of such a story will depend on the way I said 'Matt has three kids' and the immediate preceding utterances, however we rarely recall such minute details of the conversational context, so it would not be difficult to construct a plausible story on which I intended to communicate only that Matt has at least three children.

Quantifier domain restriction: It is the start of a new year and we have organised a party for the new graduate students. We have a variety of beers on offer, but there are some special craft lagers I want for myself (even though they were brought for the guests). I have stored most of the beer in the fridge, but I have put the craft lagers outside. Sally, one of the new students, arrives and asks where the beer is, so I tell her 'every beer is in the fridge'. Later on you find the craft lagers outside and ask me why I told Sally that every beer was in the fridge. In response to this challenge I might attempt to construct a story along the following lines: I had heard that Sally was a vegan, and I am aware that craft lagers often contain animal products. So when I said 'every beer is in the fridge' I didn't mean every beer we had purchased for the party, I meant every beer which was safe for Sally, as a vegan, to drink.

In these examples the speaker attempted to construct a narrative concerning their attitudes and representations of the context in which the audience's recovered meaning was not intended. In general, if recovery of an asserted content requires extensive appeal to knowledge context then the speaker will often be able to claim that the audience recovered the wrong proposition, thus disclaiming responsibility for the audience's belief. Therefore, if heavy duty appeal to knowledge context is often required for recovering what is said, and this appeal to knowledge context gives rise to the same possibilities of error to which implicature gives rise, then the deniability problem will apply to a wide range of assertions, not just implicatures.

In the examples I have given the hearers come to believe some fine grained proposition, and the speaker is able to maintain plausible deniability by claiming that a different fine grained proposition was intended. In chapter three it was noted that this is a simplification, and several other views were discussed (and rejected), including the view that speakers typically assert a very general disjunctive proposition. It is worth briefly reconsidering the disjunctive view, because it is

especially problematic for the deniability problem. The deniability problem breaks down on the view that we recover the disjunction of the propositions that the speaker could plausibly have meant. The speaker cannot plausibly deny having meant to communicate something at least as strong as the proposition recovered if the proposition recovered is the disjunction of propositions the speaker could plausibly claim to have meant.

As argued in the previous chapter, this response fails since we frequently do form testimonial beliefs in fine grained non-disjunctive propositions when there are other propositions the speaker could have claimed to have meant.   This is illustrated by the case of scalar implicature outlined above: on the disjunctive view the recovered proposition communicated would be 'Matt has either exactly three kids or at least three kids'. The problem is that this just collapses into 'Matt has at least three kids', yet it is plausible that we often take utterances like 'Matt has three kids' to communicate that Matt has exactly three children.   On the disjunctive view such readings would be rare.   A second problem which pertains specifically to the disjunctive view as a response to the denaibility problem is that the range of situations an imaginative speaker would be able conjure up in order to claim that a miscommunication has occurred will often be rather wide, and an unimaginative hearer (or just a hearer who is interpreting quickly and unreflectively) is unlikely to consider (consciously or subconsciously) the whole range of cases a speaker could construct. Thus hearers will be able to infer little from what they have been told due to the level of uncertainty over exactly what it is that they have been told. Consider the quantifier case again. I mentioned one story the speaker could conjure up in order to maintain plausible deniability, however there are many more. For example, the speaker might attempt to claim that 'beer' and 'lager' are distinct, and that the question Sally asked was asked about beer (some people do consider this to be an important distinction), or the speaker could maintain that the intended interpretation was 'every beer which is ready to drink', claiming that the beer left outside was still warm at the time of utterance. On the disjunctive view the proposition the hearer recovers would be 'every beer for the party is in the fridge, or every vegan friendly beer is in the fridge, or.....'. It is not clear why Sally would consider all these readings (especially if she is not actually a vegan), and even if she did she would not be entitled to assume that there is no other unconsidered situation the speaker could conjure up in order to claim miscommunication. Thus really all she would be entitled to take from the utterance would be that there is *some* beer in the fridge. However, we frequently take much more from such utterances. This is evidenced by the fact that Sally may well be surprised to find the craft lagers outside, or she may assume they were someone's private stash. She would not be so disposed unless she had come to believe something like 'every beer for the party is in the fridge' (she may not have explicitly

considered such a restriction, but this does not rule out forming a dispositional or unconscious belief with such a domain restriction). So I consider the disjunctive view to be implausible.

A alternative (perhaps more plausible) way of spelling out the disjunctive view holds that audiences do not take speakers to mean '*p* or *q* or...', but rather that audiences come to believe 'the speaker means *p*, or the speaker means *q*, or....'[54]. This view faces many of the same problems however, as the audience is still only able to draw a disjunctive conclusion about the subject matter of the assertion. For example, suppose I assert 'Matt has three kids', and the audience forms the belief 'Either S meant 'Matt as exactly three kids' or S meant 'Matt has at least three kids'. The audience will still only be able to draw a disjunctive conclusion about the number of children Matt has. That is, they will at best be warranted in forming the belief 'Matt has either exactly three kids or at least three kids'. This clashes with the fact that we would often form the belief that Matt has exactly three children upon hearing 'Matt has three kids'. Other problems remain too. For example, it still seems likely that the range of interpretations an imaginative speaker would be able to conjure far outstrip the number of interpretations the audience is likely to consider when forming their disjunctive view about what the speaker might have meant. Thus, although this modification of the view does perhaps seem more realistic, it still faces many of the same problems.

In the previous chapter I also considered speech act pluralism as a response along with the disjunctive view. One might think that speech act pluralism would offer a better response in this situation because it holds that speakers assert multiple propositions, and thus that they can be held to multiple propositions. In a sense this seems true. For example, looking back on Nixon's smoking gun utterance it seems clear that he could be held responsible for all the things attributed to him. Thus it does seem we undertake a commitment to many different propositions when we assert. However, it seems that speech act pluralism will not help in the cases outlined in this chapter. There are different views one might have regarding speech act pluralism and the commitments speakers undertake. One might hold that a speaker has asserted any proposition they could be interpreted as having asserted, or one could hold that speakers only assert a subset of the propositions they could be interpreted as having asserted. Additionally one could hold that a speaker undertakes a commitment to all the propositions they assert, or only to some subset. It is simply not plausible that speakers can be publicly held to all the propositions they can be interpreted as having expressed. Otherwise we would, for example, be unable to use the phrase 'Matt has three kids' to express the

---

[54] Thanks to an anonymous referee for the journal Episteme for pointing out this altervative version of the disjunctive view.

proposition that Matt has at least three kids, for we would always thereby commit to the claim that Matt has exactly three kids. This rules out the combination of the claim that speakers assert anything that they can be interpreted as having expressed and the claim that speakers undertake a commitment to all the propositions they assert. The alternative views are that   speakers assert all the propositions they can be interpreted as having expressed, but are not committed to all the propositions they assert, or that speakers are committed to all the propositions they assert but do not assert all the propositions they can be interpreted as having expressed. However, both of these views leave scope for plausible deniability. The speaker can either deny that they asserted *p*, or deny that they are committed to *p*. For example, it might be true that I happen to assert both that Matt has exactly three kids and that Matt has at least three kids whenever I utter 'Matt has three kids'. But on such a view audiences will aim not to recover what is asserted, but rather the propositions which are both intended and asserted. Otherwise communication simply would not make sense. However, we now have scope for plausible deniability again, for speakers can deny that they intended (and thus either asserted or became committed to) a particular proposition.

2. Why is the Deniability Problem a Problem?

In the previous section I outlined the deniability problem, and argued that there is no principled reason to hold that it applies only to implicature. The problem, as it applies to assertions, can be stated as follows: usually when one makes an assertion with a clear propositional content *p* one undertakes a commitment to *p*. That is, one undertakes a commitment to defend one's belief in, and assertion of *p*, or else retract the assertion. This commitment is made public in the act of assertion (see MacFarlane (2005, 2011), and Rescorla (2009) for views of assertion which place special emphasis on this fact). However, in certain cases one has another option available when challenged - one is able to deny that the audience recovered the intended proposition. Call this alternative conversational move the 'mismatch move'. When the mismatch move is available no public commitment is undertaken to defend one's assertion or provide epistemic justification for the proposition seemingly asserted.

In this section I will explain why we should be worried by the deniability problem. We should be worried about the deniability problem because the commitments speakers undertake plausibly play an important role in justifying our testimonial beliefs, and in shaping the epistemic rights we acquire when we form testimonial beliefs. The presence of the deniability problem indicates that

speakers are able to back out of certain commitments, meaning that the commitments are unable to perform their justificatory or rights shaping roles.

I start by discussing the justificatory role played by speaker commitments. I focus primarily on Fricker's own view (in order to further explicate her take on the problem) and the assurance theorists, who press the role of commitments intentionally incurred. The commitments so emphasised seem especially susceptible to plausible deniability, thus the deniability problem seems to be particularly worrying for the assurance theorist[55]. I will then briefly consider the extent to which the deniability problem might be seen to extend beyond views such as Fricker's or the assurance theorist's. Next I outline the role speaker commitments have been taken to play in shaping the epistemic rights we acquire as a result of forming testimonial beliefs. Sanford Goldberg (2006), and Benjamin McMyler (2013) have argued that the commitments speakers undertake when testifying shape the epistemic rights of the audience with respect to their testimonial belief. The presence of the deniability problem causes problems for the acquisition of these epistemic rights. Finally I argue that the presence of the deniability problem blocks one of the primary disincentives to deception.

2.1 Fricker and the Assurance Theorists.

Fricker frames her discussion in terms of what she takes to be the paradigmatic mode of transmission of knowledge via testimony - the act of telling. She argues that by telling an audience that p speakers vouch for, and take responsibility for, the truth of p. In telling someone that p the speaker presents p as being true in an act the import of which is that the hearer can form a belief in p on her say so (Fricker (2006a)). This act licences the audience to believe that p in virtue of the fact that it is the 'conventionally constituted force of her speech act' that in asserting p the speaker purports to speak from knowledge (Fricker (2006a), p594). This is a commitment in the public sphere, but it is also manifested publicly to the audience. The knowledge norm for tellings follows, Fricker thinks, from the fact that in telling someone p you offer them your word that p, and commit to it. To explain this Fricker draws an analogy with promising - it seemingly follows from the fact that in promising to perform act a you commit to doing a, that you should promise to a only if you intend to a, and $a$ing is within your power. Similarly you should not commit to p unless you know that p. It is the fact that knowledge is the norm of telling which imbues it with its epistemic force.

---

[55]   Thanks to an anonymous reviewer for encouraging me to focus on this aspect of the assurance theorist's view.

Audiences are justified in believing a speaker when they are justified in taking the speaker to be trustworthy with respect to the particular telling. The speaker's trustworthiness with respect to *p* consists in her having the following dispositional property: not easily would she assert that *p*, or vouch for the truth of *p*, unless she knew that *p* (Fricker (2006a), p 600). When you are properly justified in taking a speaker to be trustworthy, and you know they have acted in such a way that they are committed to knowing that *p*, then you are both justified in taking *p* to be true, and justified in believing that the speaker's support for *p* is sufficient to yield knowledge. Once the audience knows the speaker has asserted *p*, and is in a position to know that the speaker is trustworthy[56], then the audience is thereby in a position to know that *p*. This is how knowledge is spread via assertion according to Fricker.

So, according Fricker, speakers don't properly vouch for a proposition if they maintain plausible deniability with respect to their intention to communicate that proposition, and 'tellings' only occur when speakers do undertake such a commitment. Thus, any assertion which leaves open the mismatch move is not a telling in Fricker's sense. However, tellings (and the commitments they generate) are central to Fricker's view of testimonial knowledge. It is in virtue of the commitments tellings generate that the knowledge norm applies to them, and it is in virtue of the fact that the knowledge norm applies to them that they constitute evidence for the proposition asserted. Thus, in Fricker's framework, audiences will not be able to achieve testimonial knowledge from any assertion where the mismatch move is available (unless an alternative story about testimonial knowledge is given to supplement Fricker's 'telling' based view). If the deniability problem applies to a wide range of assertions, or if particular discourses are especially susceptible to the deniability problem, then this should sanction either a widespread scepticism, or a targeted scepticism regarding the particular discourses in question.

Fricker's view is similar in many ways to the assurance view of testimony advocated by Richard Moran (2005a), 2005b), Edward Hinchman (2005), and Angus Ross (1986). Like Fricker assurance theorists take tellings to be the paradigmatic speech act by which knowledge is transmitted, and like Fricker they take tellings to generate a special sort of commitment on behalf of the speaker. Unlike Fricker assurance theorists take the reasons to believe provided by tellings to be non-evidential (that is, the epistemic force of testimony does not primarily consist in testimony that *p* providing evidence that *p*).

---

[56] In earlier work Fricker (1994, 1995, 2006b) has maintained that in order to form a justified testimonial belief audiences must monitor speakers for trustworthiness. Presumably it is such monitoring which will grant the audience knowledge level justification of the speaker's trustworthiness.

Assurance theorists distinguish between two ways in which we may learn something by believing another's assertion. Firstly, when someone asserts *p* we may take this as evidence that they believe *p* (or perhaps gain the right to believe that they believe that p), and that, since they are likely to be a reliable belief former, we should therefore believe that p. This way of forming beliefs is, according to the assurance theorist, in principle no different from coming to believe that *p* through observing any aspect of an agent's behaviour and judging that they believe that *p*. Indeed, Moran (2005b) argues that treating a speaker's utterance as evidence about their beliefs may be worse than treating other behaviour as evidence for their beliefs due to the fact that the evidence provided by assertions is, by its very nature, doctored evidence (see Keren (2012) for a response). Assurance theorists also argue that speakers do not intend for their utterances to be treated as evidence in this way. Ross (1986) argues that in order to take the evidential view one must judge another's utterance by reference to further generalisations about their psychology and the conditions under which they are likely or unlikely to utter particular words. You need to view the assertion in a 'detached objective light, as a natural phenomenon arising from certain causes' (Ross (1986, p72)). Ross observes that we can obviously view other people's utterances this way, but thinks it is far less clear that we can view our own utterances this way.

> 'I cannot at one and the same time see it as up to me what I shall say and see my choice, as an observer equipped with a theory of speech behaviour might see it, as determined or constrained by facts about my own nature.' Ross (1986), 72.

Such an attitude would, according to Ross, be a form of disengagement from one's own actions, similar to Sartre's 'bad faith'. In order to take an utterance as it is intended by the speaker we must not treat it as evidence. The alternative, according to the assurance theorist, is to treat the speaker's taking responsibility for the audience's belief as a reason for the audience to hold that belief. The idea is that in telling the audience that *p* the speaker gives the audience permission to epistemically rely on them. When we treat the behaviour of others as evidence for their beliefs, and then form beliefs about the world on the basis of taking others to be reliable belief formers, we do not gain the ability to hold others epistemically responsible for our new beliefs. But when others tell us that *p*, and we take them at their word, we apparently do. Moran (2005a) summarises the distinction between the two ways of viewing another's testimony as follows:

> 'Corresponding to the difference between what the speaker 'gives' and what the speaker

'gives off' is the difference between what I learn from him and what I may learn from what he does and how he does it. Only in the case of what I learn from him, the person, does my relation to his belief involve the speaker assuming any responsibility for what I believe, and that makes a difference to the type of reason to believe that is obtained in the two cases.' Moran (2005a), 335.

So the speaker's public commitment to defend the audience's testimonial belief takes centre stage in the assurance theorist's view of testimonial knowledge. It is important to emphasise the role of the speaker's intentions in generating these commitments. It is by openly and intentionally communicating *p* that the speaker takes responsibility for the audience's belief. It is not clear that one can unintentionally take epistemic responsibility for another's belief. After all, consider the types of reactions we typically have to an unintentional communication of a controversial proposition. When the speaker clearly, openly, and intentionally communicates that *p* we expect them to present epistemic reasons in defence of *p*, indicating that we expect the speaker to bear epistemic responsibility for the communicated proposition. However, when a speaker unintentionally communicates *p* we do not expect them to defend the truth of *p,* rather we expect them to either defend the reasonableness of their preferred interpretation, or apologise for being unclear (we may also expect some form of compensation for any mishap which resulted from the miscommunication). This indicates that although we sometimes hold the speaker to bear some practical responsibility for propositions unintentionally communicated we do not generally hold them to bear any epistemic responsibility for such propositions. The importance of the speaker's intentions in generating these commitments is emphasised by Moran:

'Only with respect to what I have called 'personal expression', the intentional action of expressing one's belief, is the person in a position to speak for the meaning or epistemic import of what he is attesting to. With respect to whatever else may express itself in someone's speech or other expressive behaviour, while this may indeed be a source of knowledge for the audience, they are on their own as far as assessing its epistemic significance goes. Since beliefs which are revealed in these ways need not even be known by the speaker himself, the hearer (or observer) cannot assume that the speaker is in a position to offer support or justification for what may be garnered in this way, nor that he speaks with any authority about the meaning or general significance of the belief which manifests itself in his speech or other behaviour.' Moran (2005a), 342.

As we observed when discussing Fricker's view, the speaker's public commitment to defend the audience's belief is precisely the type of commitment the deniability problem undermines. This is worrying if the mismatch move is widely available, or if there are particular discourses in which it is widely available. In cases where audiences have to rely heavily on knowledge context the audience's attitude toward the speaker seems to be no less one of trust in the speaker than in the assurance theorist's paradigm cases. Therefore, if the assurance theorist is correct that in the paradigm cases we do not treat assertions as evidence then there are a range of utterances which we do not typically treat as evidence, but which can only provide epistemic reasons in virtue of being treated as evidence (since speakers don't undertake the assurance theorist's required type of commitment in making these utterances). Once again this sanctions a scepticism about the discourses in question.

Indeed, it seems that the assurance theorist's emphasis on the speaker's intentions makes the problem especially worrying. This is because the story the speaker must tell in order to maintain plausible deniability about their own intentions will usually require appeal to little more than claims about their own internal mental states (over which they have epistemic authority). In fact a speaker could even maintain that they were simply not paying attention to what they were saying in order to maintain that a rather obvious interpretation was unintended. In such cases we would certainly hold a speaker responsible for some wrong doing, however we would often still accept that they did not intend to communicate what they in fact did communicate. Thus, if the assurance theorist is correct about the role of speaker intentions in generating the types of commitment required for testimonial justification then it will often be fairly easy for speakers to make the mismatch move, meaning that it will be widely available.

Of course, there are many commitments and responsibilities we undertake when asserting over which we do not have intentional control. For example, we do not have intentional control over which propositions will actually be communicated by our utterances. However, we still have a moral duty not to make an assertion if it is predictable that in doing so we will communicate something false[57]. If we do make an utterance which predictably results in a false belief then we will often be held responsible for the resultant belief. This responsibility is not epistemic, even if we believe the predictably communicated proposition we are not thereby duty bound to defend its truth. Nonetheless, we will often be held practically or morally responsible for any misfortune which results from the false belief. It may be thought that these commitments are sufficient to

---

[57]   Thanks to an anonymous reviewer for pointing this out.

provide a basis for testimonial justification. We might reason, in a similar way to Fricker, that since we are committed to not predictably communicating falsities we will try to only predictably communicate truths. Thus, if a speaker makes an assertion such that it is predictable that p would be communicated by that assertion this provides evidence that the speaker knows that p. Such a view would not provide the kind off anti-reductionist justification the assurance theorist seeks. However, speakers have far less plausible deniability over what will be predictably communicated by an utterance, so perhaps such a view could at least provide a reductionist account of the role of speaker commitments in generating testimonial justification whilst mainly avoiding the deniability problem.

Ultimately I think that this view contains more than a grain of truth. However, it is not without its problems. Firstly, speakers are generally far more careless about what they might predictably communicate than what they clearly and openly communicate. Thus, it is not clear just how often audiences will be justified in believing that the speaker knows that p on the basis of the fact that the speaker made an utterance such that it is predictable that that utterance would communicate p. This will depend on just how easily predictable the audience's interpretation was, and on the context (for example, speakers are likely to be far more careful about what they might communicate in a high stakes context than a low sakes context). Secondly, although speakers will have less plausible deniability about what was predictably communicated by their utterance, they will still have some. That is, they might have some plausible deniability regarding their own ability to predict the resultant interpretation. This will involve the construction of a story about the speaker's representation of the context at the time of utterance such that a reasonable agent with that representation could make their utterance without being able to easily predict that they would communicate p. Even if such a story fails to establish that the speaker was not in a position to predict the resultant interpretation it might still convince the audience that their own interpretation was less obvious to the speaker than they thought, thus making the speaker less criticisable for their utterance. Thus it is unclear to what extent such a view does avoid the deniability problem.

2.2 Deniability and Buck-Passing.

In the previous section I argued that the deniability problem is worrying for those who emphasize the role of speaker commitments in testimonial justification. An agent's right to form a testimonial belief is, according to such views, at least partly grounded in the fact that speakers take

responsibility for the audience's belief. However, some authors have argued that the epistemic role of speaker commitments extends beyond the justification of testimonial beliefs. They argue that testimonial knowledge comes with distinctive epistemic rights which sets it apart from other forms of knowledge. These rights pertain not to our acquisition of testimonial beliefs but rather our retention of them in the face of challenges. As Goldberg (2006) emphasises this illustrates a sense in which whatever one says about testimonial justification (that is, whether one is a reductionist or an anti-reductionist about testimonial justification) testimonial knowledge is epistemically distinctive.

Both Goldberg (2006, 2011) and McMyler (2013), when discussing what is distinctive about testimonial knowledge, have emphasised the fact that audiences have the right to 'pass the buck' in response to challenges. That is, when one forms a testimonial belief on the basis of someone else's say so one has the right, when challenged to retain one's belief and defer to the original testifier. For example, imagine that Sammy tells Lizzie that oats lower cholesterol, and Lizzie then tells Mark that oats lower cholesterol. Mark is sceptical and challenges Lizzie's assertion. At first Lizzie says something vague about low density fibre and bacteria in the intestines, but Mark is not convinced and continues to challenge Lizzie's assertion. At this point Lizzie is able to retain her belief but defer to Sammy, saying 'Well Sammy told me, ask her about it'. That is, Lizzie is able to pass on the responsibility for defending her belief to Sammy, since Sammy was the original testifier. The right to pass the buck in this way is a downstream epistemic right which agents acquire through forming testimonial beliefs. Other belief forming methods (such as perception) do not generate such a right. Thus, the fact that such rights are acquired seems to be a distinctive epistemic feature of testimonial beliefs. Goldberg puts the point as follows:

> 'My main contention is that testimonial knowledge is a distinctive kind of knowledge in that this sort of knowledge, but no other, is associated with a characteristic expansion in the sorts of epistemically relevant moves that can be made by the subject in her attempt to identify the direct epistemic support enjoyed by her belief.' Goldberg (2006), 133-134.

> '...this feature of testimonial knowledge reflects the fact that there is something epistemologically distinctive about relying on the epistemic authority of another rational being: it is because of what is distinctive in relying on the epistemic authority of another rational being, that there is a characteristic expansion in the sort of moves that can be made

in defence of a belief acquired on such authority. The characteristic expansion, I suggest, is that testimonial knowledge gives rise to the hearer's right to *pass the epistemic buck* after her own justificatory resources have been exhausted.' (ibid: 134)

As emphasised by Goldberg (2006), one can recognise this no matter what one's view of the original justification for testimonial beliefs. This is because the rights pertain not to the conditions under which it is acceptable to form a testimonial belief, but rather the moves one can make when those beliefs are challenged. Moreover, Goldberg (2011) later argues that these distinctive epistemic rights derive from rather uncontroversial features of testimony and assertion. Assuming that assertion has an epistemic norm (that is, assuming that the propriety of an assertion that *p* requires that the speaker be in some way positively epistemically situated with respect to *p*) the act of asserting creates mutual knowledge amongst the speaker and audience that the speaker has performed an action the propriety of which requires that they are epistemically well situated with respect to *p*. Thus, when challenged the audience knows that there is a further body of information (that possessed by the speaker) to which they can appeal in defence of their belief[58]. As emphasised by both Goldberg and McMyler these rights are genuinely epistemic since they pertain to the ways in which agents are answerable for their beliefs (that is, the epistemic responsibilities they bear with respect to those beliefs). These are taken to be genuinely epistemic features of testimonial knowledge which are not shared by other forms of knowledge. McMyler provides a succinct statement of his reasoning which is worth quoting in full:

> '..the responsibilities involved in epistemic buck passing are genuinely epistemic responsibilities, responsibilities that pertain to the way in which cognitive agents are distinctively answerable for their beliefs. Belief, I take it, is a commitment-constituted attitude. To believe that *p* is to commit oneself to a positive answer to the question whether p. A believer is thus answerable for being so committed. She is open to criticism that bears on the content of her commitment, criticism that bears on the question whether p. Plausibly, one aspect of the way in which subjects are thus answerable for their beliefs concerns their epistemic conduct in the face of reasonable challenges to their beliefs, where reasonable challenges to their beliefs involve the presentation of evidence that counts against their beliefs. Typically, when confronted with such a challenge, a rational epistemic agent ought to either find some way to meet the challenge—some basis upon which to rationally

---

[58] However, it is unclear to me how these facts alone generate any form of duty in the speaker to defend the audience's belief, and thus where the audiences right of deferral derives from.

discount the evidence presented—or else give up her belief. When it comes to beliefs that are based on being told something by a speaker, however, an epistemic agent is entitled to maintain her belief without meeting the challenge herself by instead passing the epistemic buck back to the speaker. If an audience comes to believe that p on the basis of a speaker's telling, and if a third party challenges the audience's belief by producing evidence that counts against p, the audience is entitled to defer responsibility for meeting the challenge back to the original speaker, whereupon the original speaker is epistemically responsible for meeting the challenge.' McMyler (2013), 1067-1068.

I take these considerations to establish that, at least in some cases, testimony generates the right to pass the epistemic buck, and that this is a genuine epistemic right. In such cases the believer will be less responsible for their belief than they would otherwise be. Moreover, this right to defer to the someone else in defending one's belief illustrates a sense in which such beliefs, and the rights we have regarding them, are distinctly social. This brings us to another sense in which the deniability problem seems genuinely problematic. The deniability problem undermines these speaker commitments. This is particularly clear on Goldberg's (2011) statement of the view which rests on there being mutual knowledge between the speaker and hearer that the speaker has performed an action the propriety of which requires that they are epistemically well situated to a particular proposition p. When the mismatch move is available the speaker is able to provide a defeater to this mutual knowledge claim, meaning that they can back out of their commitments. If speakers undertake no public obligation to defend the audience's belief then audiences don't gain the right to defer to the speaker when challenged. Thus, when the mismatch move is available audiences will form testimonial beliefs but will not acquire the distinctive epistemic rights which standardly accompany such beliefs. If there are discourses which are particularly susceptible to the deniability problem then audiences will bear more individual responsibility for the beliefs acquired from such discourse than they think they do (which could lead to overly credulous belief forming).

2.3. Plausible Deniability and Testimony Policing.

So far I have outlined two reasons to worry about the deniability problem. First, it undermines testimonial knowledge on telling based views. Secondly, when the mismatch move is available audiences will not gain the epistemic rights distinctive of testimonial beliefs, meaning that they will bear more responsibility for their beliefs than they think they do. There is one further reason to

worry about the deniability problem. Usually when someone falsely asserts that *p* we can criticise their assertion, and criticise them as an asserter. People who are caught asserting falsehoods are publicly labelled as liars. This is a strong normative assessment, and the chance of being so labelled provides a disincentive against asserting falsehoods. However, when the mismatch move is available it is far harder to make such an accusation. The speaker can always respond claiming that they never intended to communicate the falsehood in question. Thus, if there are discourses which are particularly susceptible to the deniability problem speakers in such discourses will have far less incentive to speak honestly (Goldberg (2013) presses a similar line in order to explain why we should be sceptical of anonymous internet testimony). Of course, we do have other normative assessments available. We can accuse a speaker of being misleading. However, this accusation carries far less normative force, especially if we cannot establish that the speaker has been intentionally misleading. And to establish such a thing will usually require a great deal of information about, for example, the speaker's motivations and their knowledge of their audience.

This concludes the first half of the chapter. I have outlined the deniability problem, argued that there is no principled reason to hold that it applies only to implicatures, and provided three reasons to consider it genuinely problematic. Firstly, it undermines testimonial knowledge if one embraces a telling based view of testimony. Secondly, it undermines the distinctive epistemic rights typically acquired via testimonial belief formation. And finally it removes the disincentive for speakers to be intentionally misleading with their assertions. In the second half of this chapter my task is to identify the scope of the deniability problem for assertions.

3. The Contextualist Puzzle.

I have outlined the deniability problem, arguing that there is no principled reason to hold that it applies only to implicatures, and I have provided three reasons to consider it genuinely problematic. What I have not done is argued that a significant number of assertions actually suffer the deniability problem (although it should be clear from the argument of chapter three that it will apply to at least as many cases as the recovery problem). It is to that task I now turn. I start by considering a puzzle arising from the recent debate over context sensitivity in natural language, which seemingly suggests that the deniability problem is extremely widespread (far more widespread than we would intuitively think). I then discuss some responses to the puzzle which allow us to narrow the scope

of the problem back down. We end up with a loose set of criteria for identifying discourses which will be particularly susceptible to the problem.

The first part of the puzzle is the seemingly widespread context sensitivity of natural language. The list of context sensitive uses of language includes indexicals, demonstratives, gradable adjectives, comparative adjectives, definite descriptions, indefinite descriptions, adverbs, conditionals, modals, quantifiers, predicates of personal taste, possessives, incomplete adjectives, psychological attributions, moral attributions, perhaps knowledge ascriptions, non-sentential assertion, vagueness, metaphor, hyperbole, and loose talk. This constitutes a large portion of our language use. Some theorists (radical contextualists) go even further, holding that most, perhaps even all, of our language use is context sensitive. Such theorists note, for example, that seemingly context insensitive terms can be used in many different (and incompatible) ways in different contexts (see, for example, Travis (1985), Bezuidenhout (2002)). They also provide general theories of utterance comprehension which entail that we always engage in pragmatic processing to recover what is said. On such views, even when the recovered proposition is the proposition determined by the literal meaning of the terms used, such processing is employed to determine that it is the literal meaning (if there is such a thing (see Recanati (2004))) rather than some alternative 'modulation' which is most appropriate.

So, it seems that there is a lot of context sensitivity in natural language, especially if you believe the radical contextualist. This fact by itself shouldn't worry us. What is worrying is that the mechanisms by which we resolve context sensitivity arguably have the same 'ambiguous epistemics' that Fricker argues give rise to the deniability problem for implicatures. It has already been argued (in chapter three) that these ambiguous epistemics give rise to epistemic problems. Cappelen and Lepore (2004) go even further, arguing that the mechanisms which contextualists postulate to explain context sensitive communication would seemingly make communication miraculous. We have already noted that audiences have to quickly appeal to a wide range of knowledge in order to recover the content of utterances. Cappelen and Lepore point out that if we have to rely on all this information just to recover the content of simple context sensitive utterances such as 'philosophy is fun' then it seems miraculous that we generally tend to communicate smoothly and successfully with context sensitive terms. The fact that we have to rely on such knowledge means that there are many ways the recovery of a given proposition can go wrong. This, as we saw, was precisely what generated the deniability problem for implicatures (and the recovery problem). There were many ways the recovery of an implicature could go wrong, thus the speaker could easily claim that

something did indeed go wrong, disclaiming responsibility for the audience's belief. To get an idea of some of the ways in which we rely on knowledge context to recover what is said consider the two leading approaches to context sensitivity.

> Saturation: For many sentences the phonetically articulated elements don't exhaust the syntactic structure, there are non-phonetically articulated syntactic elements which need to be assigned values in order for a complete proposition to be expressed. This is usually taken to consist in there being hidden variables in the underlying logical form which take particular types of value[59]. Some variables will be bound by linguistic material from earlier on in the same sentence, meaning that their values will be easily recoverable. Indeed, one of the primary reasons for positing hidden variables in clauses such as 'it is raining' is that it looks as if there are longer sentences in which they have a bound reading. However, when such sentences are not embedded the audience must employ knowledge context in order to work out the intended value of the variable (Stanley (2000) pushes this line of reasoning).

Often the values which need to be assigned are complex and fine grained. For example, they include properties to restrict quantifiers, and comparison classes for gradable adjectives. The subtle differences in possible values, and the sensitivity of these values to small changes in the mutual goals and presuppositions of the speaker and hearer can, in many cases, give the imaginative speaker scope to make the mismatch move. Take, for example, quantifiers - in many cases one can claim the intended range of a given quantifier to be a restriction on the range the audience attributes. You merely need to be able to identify a subset of the domain the audience attributes such that the members of that subset have some distinguishing feature, and identify an aim relative to which this feature would be relevant such that you could, with at least some level of plausibility, have taken it to be mutually presupposed that it was a conversational aim. For example, in the quantifier case in section one the distinguishing feature of the selected subset was that it contained only vegan friendly beers, and the mutual goal was to identify beers which Sally would be able to drink. Speakers are frequently slippery with what they say in precisely this way.

> Modulation: In cases of modulation a constituent has its meaning adjusted, and thus contributes something new to the truth conditions of the utterance. The concept we end up

---

[59] Not all theorists take saturation to be mandated by variables in the underlying logical form. Some theorists deny the existence of such variables, yet maintain that utterances of the sentences in question fail to express full propositions, arguing that audiences supply unarticulated constituents in order to fill in the gaps by appeal to the knowledge contexts (Carston (2002a), Hall (2008), Perry (2001), Recanati (2004)).

with will be related to the concept encoded in the constituent before modulation takes place, but will usually serve the speaker's purposes better. For example, a concept may be narrowed and thereby assigned a new meaning which applies to a subset of the original extension (e.g. when Sally, who has just left her timid and underachieving boyfriend says 'I need a man', the concept encoded by 'man' is narrowed in order to apply only to men with particular features commonly associated with masculinity (Carston (2002a)). A concept may also be loosened to generate a new concept which applies to more things than the original concept. Modulation occurs in response to the audience's perception of the demands of the context. It is controversial how the process of modulation works. Some (e.g. Recanati (2004)) think that potential meanings vie for cognitive activation with the most salient meaning being assigned. Others (e.g. Carston (2002a, 2002b), Carson and Wilson (2007)) think that potential meanings are ranked in order of salience and assessed for relevance (a balance of cognitive effects and cognitive effort) until an expectation of optimal relevance is met, at which point the meaning which meets the expectation is assigned.

It should be clear from this that modulation also relies heavily on knowledge context, that the type of context sensitivity it accounts for is common (consider how mundane most of the examples were), and that there will often be many ways an audience can go wrong in recovering the correct proposition. When discussing saturation it was noted that in cases where the values assigned are complex the precise values assigned will be very dependent on knowledge context, thus creating a lot of scope for genuine miscommunication, and the mismatch move. The same is true of enrichment. Indeed, since enrichment is even less constrained it seems the problem will be even worse. Enrichment occurs only in response to the audience's impression of the demands of the context, and the values involved are less constrained than in cases of saturation, giving even further scope for the speaker to claim mismatch. Additionally, speakers can claim to have intended to be interpreted more or less literally than they were (an option which is not obviously available in cases of saturation). All in all, it seems that the problems which arise for saturation not only arise, but are multiplied in cases of enrichment[60].

---

[60] I think that many contextualists (especially those who focus on modulation and free enrichment) recognise the fact that their theories involve the sorts of 'problematic epistemics' pointed to by Fricker. Many of these theorists have worried about related problems which their own theories raise for communicative success. This has led many prominent theorists (such as Bezuidenhout (1997), Carston (2002), Heck (2002), Recanati (2004), and Sperber and Wilson (1986)) to conclude that communicative success does not require the precise sharing of contents between speaker and audience, but rather entertainment of similar propositions. They endorse this view partly as a result of their recognition of the epistemic difficulties which arise in the recovery of propositional contents. If communicative success required that speakers and audiences shared identical contents then communicative success would be rare due to the epistemic difficulties involved in recovering a propositional content identical to that intended by the speaker (many of these theorists also raise worries about the sharing of Fregean contents).

We now have the two components of the puzzle generated by contextualism. Firstly there is reason to believe that there is a great deal of context sensitivity in natural language, secondly there is reason to think that this context sensitivity has the same problematic epistemics which gave rise to the deniability problem for implicatures. This suggests that the deniability problem is very widespread. This is a radical conclusion, and one we should be eager to avoid, since it seems quite obvious that speakers don't have plausible deniability about what they say the majority of the time.

4. Responding to the Contextualist Puzzle.

In this section I consider a series of responses to Cappelen and Lepore[61]. I will argue that these responses fail as general responses to the deniability problem. Nonetheless, they are worth considering because they illustrate a sense in which it may be harder for speakers to make the mismatch move in certain contexts. We will see that a discourse must have certain features in order to block plausible deniability for context sensitive assertions. Thus we will be able to identify discourses which are particularly susceptible to the deniability problem.

Ishani Maitra (2007) has argued that certain contextual values are more natural and frequently applied than others, and that we will assign such values unless we have reason not to. For example, if I say 'elephants are big' it will be more natural for you to assign the comparison class of species rather than some other comparison class, for example the class of large mammals. You would generally only assess the claim relative to the class of large mammals in response to additional information which made that reading more likely. If this is the case then speakers and hearers will usually converge on the same contents because audiences will not need to appeal to knowledge context to a problematic extent.

I think it should be clear that this response won't solve our problem. Even if it does help to explain how interlocutors converge on contents it still leaves a lot of scope for speakers to detach themselves from the proposition recovered. If the audience settles on the more common or natural reading then the speaker may claim that further information should have been appealed to in

---

[61] Cappelen and Lepore's own position is that we must separate semantic content and speech act content, and that speech act content but not semantic content is context sensitive. Many of the same problems arise for their theory as arise for the contextualist, because the proposition we will form a testimonial beliefs in will usually not be the minimal semantic content, but rather the context sensitive speech act content.

interpretation. Likewise, if the audience moves away from the more common reading the speaker can claim that the audience drew too much from context and that the common reading was intended. This is not to say that Maitra's response is without use. Certainly in some contexts, and for certain common conversational tasks, standard meanings for terms emerge. And certainly in some contexts there will be a meaning so obviously more natural than all the others that the speaker has no scope for plausible denial concerning what they meant. But it is far from clear that the majority of cases are like this.

A related but more promising response has been provided by François Recanati (2010). Recanati doesn't postulate a set of natural or common meanings, however he does postulate that there are important psychological commonalities which dispose people to converge on the same meanings of terms, and which enable us to recognize how others intend concepts to be modulated. For example, people are disposed to recognise the same similarities between the situation of application for a concept, and other situations to which the concept does not straightforwardly apply. As a result they are able to extend or narrow the use of term a which expresses that concept, perhaps modulating the concept along these dimensions of similarity to apply to the new case (see also Bolinger (1968)). Our interpretation is also taken to be guided by sets of implicit biases which are common across speakers and audiences. Recanati does not provide any examples of such biases. However he does point to some biases postulated by psychologists working on the early acquisition of lexical meanings. For example, it is argued by Bloom (2000) that early acquisition of lexical meaning is guided by a whole object bias (a bias toward taking a whole object rather than the parts of an object to be the referent of a term). It is conceivable that a network of such common biases guides our ordinary interpretation and helps us assign values to context sensitive terms[62].

This does not solve the deniability problem. Firstly, this network of biases and abilities may often lead audiences in the right direction, but (especially if such biases are rooted in theory of mind) they will still be dependent on assumptions about the context to which the speaker can appeal to in order to claim miscommunication. Thus the mismatch move will still be available. Secondly, this response has a rather narrow scope. Although there may be similarities, or types of similarity, which humans as a kind are more disposed to recognise, and biases toward objects which can be categorised in certain ways, these are extremely unlikely to exhaust the range of similarity judgements and psychological mechanisms which guide interpretation. It would be very surprising

---

[62] I suspect that this response only secures convergence on similar rather than identical propositions. This shouldn't worry Recanati since he maintains that similarity of content is sufficient for communicative success (as do Bezuidenhout (1997), Carston (2002), and Heck (2002))).

if we were not also guided by similarity judgements and biases which are moulded by our individual experiences. This seems especially true when we are dealing with more abstract concepts. This creates a greater chance of genuine mismatch, and more scope for speakers to make the mismatch move.

A third response draws our attention to communication as a collaborative affair. It has been argued that conversational participants don't allow situations to arise where there is any realistic chance of content mismatch (Perrini (2009), Recanati (2010)). This is because speakers don't just make an utterance, get interpreted, and move on. Rather, there are collaborative checks in place to ensure understanding. Both speakers and audiences track each other's facial expressions, tone, and body language for signs of misunderstanding or mismatch. Additionally, if there is uncertainty about what was said the audience asks for clarification (Clark and Krych (2001)).

In responding to the deniability problem the thought would be this: both audience and speaker collaborate to establish a shared meaning, and audiences can refer back to this when a speaker attempts the mismatch move, thereby blocking plausible deniability. This response also fails. It is only in cases where either the audience is aware of their lack of understanding, or when their subsequent interactions with the local environment indicate misunderstanding, that mistakes are corrected. When the speaker and hearer are not coordinating on a mutual task involving the local macro level environment the hearer will have nothing to refer to when calling out the speaker. Additionally, many similar but epistemically distinct contextual values (e.g. quantifier domain restrictions) will have very similar behavioural consequences, meaning that misunderstanding won't immediately generate behavioural evidence of miscommunication. So the response is somewhat limited in scope.

So far I have surveyed a series of responses to Cappelen and Lepore and found them lacking as responses to the deniability problem. Nonetheless, I think they can teach us something important. These responses draw our attention to a set of resources to which an audience can appeal in certain contexts in an attempt to call out a speaker who is attempting the mismatch move, thereby blocking plausible deniability. For example, if there is a clear common use for a term (or common default contextual value), and the audience reasonably assigns such a value only to be met by the mismatch move later on, then the audience is able to maintain that the speaker should have been more explicit about their intention, maintaining that they are partly responsible for the resultant belief. The same goes for modulation based on similarity relations and biases. This is especially true in cases where

there are checks in place related to some mutual task. If someone acts on the basis of their understanding, and at that point the speaker fails to flag any misunderstanding, then this goes some way to confirming the audience's initial interpretation, making it far harder for the speaker to make the mismatch move without descending into absurdity. In general, the more the audience is able to check that they have the correct understanding, and the more obvious the default understandings are, the harder it will be for a speaker to claim mismatch without absurdity.

These resources are rather limited, and in one-off instances they may often prove ineffective. However, if speakers repeatedly try to employ the mismatch move in order to avoid commitment in circumstances where these resources are available then audiences will be able to call them out on their frequent misleading behaviour. It may be plausible that in a one-off case the speaker intended the audience to assign a more esoteric meaning to a term than they did. However, it becomes far less plausible in a long run of cases. Repeat offenders will lose plausible deniability. Moreover, speakers have motive to avoid appealing to the mismatch move in contexts where checks are in place, for if they make the move frequently then they will quickly lose credibility as an informant.

The deniability problem now seems somewhat less worrying. However, it was only weakened for discourses where audiences have the resources to call out the speaker by appeal to standard meanings, very obvious ways of extending a meaning, or checks which serve to reliably confirm understanding. It is a partially empirical question how many discourses actually have these features. However, I think it is likely that some important discourses lack them. The deniability problem still arises with its full force for such discourses. These will be discourses in which context sensitivity (especially more unconstrained context sensitivity such as modulation) is rife, which don't involve coordination on macro level tasks, where the values or modulated concepts are complex or abstract, and where there are no highly standardised or clearly stated contextual values. Such discourses provide speakers with a lot of scope to make the mismatch move without losing much credibility. There will be more resources to which speakers can appeal in order to claim misunderstanding, and fewer checks an audience can appeal to in order to call them out or hold them responsible. It is not the task of this chapter to establish conclusively that any particular discourse has such features, however several important discourses do seem to be candidates. For example, religious discourse is arguably rife with context sensitivity (consider the many and varied religious conceptions of salvation, love, and even God (see Alston (2005), and Scott (2005) for useful overviews discussing the context sensitivity of religious language), it has a highly abstract subject matter, and it is not clear that there are always sufficient efforts put in place to coordinate on

precise explicit meanings. Likewise much public political discourse seemingly has such features (for example, consider the possible varied meanings of terms like 'class warfare'). Indeed, this no doubt adds to the stereotype of politicians as slippery and dishonest. Another important candidate seems to be ethical discourse outside of academic settings (where there are often norms which require precision and coordination on standard meanings). Such discourse is abstract, arguably context sensitive, and seemingly lacks a norm requiring explicit joint efforts to coordinate on precise meanings. These worries are strengthened by the studies discussed by Keysar, Barr, and Horton (1998) which appear to show that audiences initially interpret egocentrically (on the basis of their own private set of assumptions rather than the common ground), and only revise these interpretations when they know that their own set of assumptions differs from that of the speaker in such a way that it would generate miscommunication. This illustrates that as long as there is the illusion of understanding audiences will not check for coordination with the speaker, suggesting that there could be a significant amount of actual miscommunication which gets missed. Indeed, in many cases, such as the sermon from the pulpit, or perhaps a political speech, the speaker's assertion and intended modulations may be backed by an understanding of theology, or politics which far outstrips that of their audience. The educated priest's conception of god may be quite different from that of their congregation. These different factors effecting the way speakers and audiences modulate the relevant concepts are also likely to generate actual mismatch. And wherever there is scope for actual mismatch the mismatch move is available (and gains additional plausibility). Of course, it is beyond the scope of this chapter to establish that any of these discourses do have the features in question. However, given the importance of such discourses in our everyday lives this seems like a worthy question for further research.

5. Conclusion.

I have outlined the deniability problem for assertions, explained why it is genuinely problematic, and presented a line of reasoning which seems to suggest that the problem is very widespread. I then looked at several ways of narrowing the scope of the problem back down. It was found that the deniability problem would be less problematic in discourses with certain features. I suggested a selection of important discourses which may still face the problem.

It is not clear how we should react once we discover that a discourse faces the deniability problem. I suspect that the correct reaction will vary between different discourses. One reaction may be to

try and establish a set of precise meanings within the discourse, and eliminate context sensitivity as much as possible (for example, if certain areas of academic discourse were found to face the deniability problem, this   would probably be the more appropriate response).   An alternative response would be to give less weight to testimony in the problematic discourse. One could treat knowledge regarding the subject matter as necessarily personal rather than social. One could re-conceptualise the role of apparent testimony in the discourse, perhaps taking it to be expressive, or seeing it as intended not to bring about belief but rather reflection or some other attitude. And another alternative would just be to view the discourse with scepticism. The plausibility of any given response will depend on the discourse in question. Finally, one might simply choose to weaken the focus on speaker commitments in one's theory of testimony[63].

Chapter Five: Epistemic Injustice in Utterance Interpretation.

0. Introduction.

Miranda Fricker (2007) explores a type of injustice which arises when the credibility judgements we make about speakers are informed by prejudicial stereotypes. In this chapter I argue that by focusing only on credibility judgements we miss an important species of injustice which arises earlier in the process of testimonial belief formation. The same prejudicial stereotypes which infect credibility judgements can also infect our interpretation of the speaker, leading to uncharitable miscommunication (call the injustice arising from prejudicial interpretation 'interpretative injustice'). There are several harms caused by interpretative injustice. Firstly, as in the case of lowered credibility judgements, interpretative injustice prevents certain groups from being able to efficiently communicate knowledge to other (perhaps more powerful) groups. Secondly it results in speakers being held epistemically responsible for propositions they never intended to communicate. And thirdly, it contributes to the illusion that prejudicial low credibility judgements are epistemically justified. I close by arguing that if Fricker's strategy for treating epistemic injustice is implemented in absence of a treatment of interpretative injustice then we risk epistemically harming the hearer with little benefit to the speaker. Thus epistemic injustice and interpretative injustice are best treated in tandem.

---

1. Epistemic Injustice.

When we learn from testimony we usually assess the testimony for credibility. This includes an assessment of the over all credibility of the speaker with respect to the topic of their statement, as well as other factors such as the plausibility of the content of the statement, and circumstantial factors which may make the speaker less reliable in a particular context. Generally, when we assign the testimony a high degree of credibility we believe it (all else being equal), when we assign it a low degree of credibility we don't. Such assessments are usually spontaneous, and reliant on general heuristics that we apply (perhaps sub personally) in particular situations. For example, if someone appears drunk then we may instinctively assign them a low credibility. This is based on the stereotype that drunk people are poor informants, or an association between the concept of drunkenness and other concepts we associate with poor informants (for example dishonesty)[64]. Such reliance on general stereotypes or associations seems essential to our ability to quickly and reliably make credibility judgements. However, this reliance on stereotypes or associations can also lead us astray. Historically there have been many stereotypes linking disadvantaged groups to factors which we associate with low credibility. For example, there was (and to some extent still is) a common stereotype of women as being emotional, irrational, overly intuitive, and naive. If such stereotypes are salient to a hearer then they may inform the hearer's credibility judgements (even if the hearer does not explicitly endorse the stereotype). The result is that the hearer will assign women low credibility as informants, at least with respect to certain topics.

Fricker points to several harms which result from epistemic injustice. Firstly, when we assign someone low credibility on the basis of factors such as race or gender we wrong them as a knower. Fricker argues that the capacity to know and to give knowledge is a fundamental human value. Thus, in limiting some individual or group's ability to give knowledge we thereby restrict their ability to engage in an activity fundamental to human value. As Fricker (2007) puts it, "the capacity

---

[64] Fricker frames her discussion in terms of stereotypes, and I will mostly follow her in this. However, I think that the notion that we appeal to stereotypes (most easily conceived of as a type of generalisation) in credibility judgements or interpretation over-intellectualises these processes. I think it is more likely that we associate certain concepts with others, and that the salience of one concept (for example intoxication) may thereby raise the salience of other concepts (for example dishonesty), and that these associations guide our interpretation and credibility judgements. Likewise, if someone harbours anti-Semitic biases the concept 'Jew' might raise the salience of 'greed', which could in turn influence their credibility judgement when it comes to testimony relating to money. Here the agent is not appealing to some sort of anti-Semitic generalisation like 'Jews are greedy' (they might not even consciously endorse such a generalisation), it is the mere association of Jewishness with greed which does the job, leading to epistemic injustice.

to give knowledge to others is one side of that many-sided capacity so significant in human beings: namely, the capacity for reason" p 44. By assigning someone a low credibility on the basis of prejudicial stereotypes we prevent them from giving knowledge, and thereby exclude them from the 'community of epistemic trust'. Fricker even contends (drawing on Bernard Williams) that the ability to share one's beliefs plays an important role in the psychological development of one's social identity through the steadying and securing of one's beliefs. When limited in our ability to convey knowledge we are, according to Fricker, limited in our ability to develop our own social identity.

Fricker also identifies a variety of secondary harms which arise from prejudicial credibility judgements. Our ability to share information is essential for our ability to partake in a large number of social activities which are central to the good life. For example, one's ability to progress in one's career, or to effectively defend oneself from false allegations[65] will usually depend on one's ability to effectively communicate one's knowledge. Indeed, usually when we try to impart knowledge we are doing so for some particular reason, in service of some more general task. When one is restricted in one's ability to share knowledge one is thereby restricted in one's ability to partake in any task in service of which one may wish to impart knowledge. Epistemic injustice also causes harm in the sense that it can lead people to lower their opinion of their own epistemic abilities. If you are treated as unintelligent for a long enough period you will eventually start to believe it yourself. If you are treated as unreliable on a particular topic you may lose confidence in your beliefs on that topic, and thereby lower your credence bellow the threshold required for knowledge.

An ideal agent would instinctively appeal only to non-prejudicial stereotypes when assessing a speaker for credibility. However, most of us are not ideal in this respect. Something must be done to remedy this situation. Fricker's proposed solution is that we attempt to develop the virtue of testimonial justice. This is a corrective virtue whereby we actively reflect on the prejudices we may harbour and develop a sensitivity for the kinds of situation in which we may be led to unjust credibility judgements as a result of these biases. In such situations we should not rely on our quick intuitive judgements, but should try to make an accurate and unprejudiced judgement. Fricker puts the point as follows:

'When the hearer suspects prejudice in her credibility judgement—whether through sensing

---

[65] One of Fricker's key examples is Harper Lee's (1960) novel 'To Kill a Mocking Bird', in which Tom Robinson, who has been falsely accused of rape, is unable to defend himself as a result of the racial prejudices of the time preventing his testimony from being trusted.

cognitive dissonance between her perception, beliefs, and emotional responses, or whether through self-conscious reflection—she should shift intellectual gear out of spontaneous, unreflective mode and into active critical reflection in order to identify how far the suspected prejudice has influenced her judgement. If she finds that the low credibility judgement she has made of a speaker is due in part to prejudice, then she can correct this by revising the credibility upwards to compensate. There can be no algorithm for her to use in determining how much it should be revised upwards, but there is a clear guiding ideal. The guiding ideal is to neutralize any negative impact of prejudice in one's credibility judgements by compensating upwards to reach the degree of credibility that would have been given were it not for the prejudice.' Fricker (2007), 91-92.

The hope is that eventually such reflective adjustment will condition us in such a way that our automatic intuitive credibility judgements become unprejudiced.

2. Interpretative Injustice.

Fricker's brand of epistemic injustice arises as a result of the way in which our credibility judgements are guided by potentially prejudicial stereotypes. However, credibility judgements occur late in the process of testimonial belief formation. Other stages in this process are likewise guided by a reliance on stereotypes. In particular, our interpretation of the utterances themselves will often be guided by stereotypes[66]. This is especially true in cases involving context sensitivity, loose talk, unfamiliar dialects or accents, noisy environments, implicature etc. In such situations we must appeal to our knowledge of the context, including what we know about the speaker (their likely goals, interests, beliefs, background, intelligence etc.) in order to reach a verdict on what proposition they are intending to communicate. These judgements, generally being fast and sub personal (like credibility judgements), will also rely on stereotypes and associations.

Interpretative injustice is the phenomenon whereby a hearer's employment of prejudicial stereotypes results in the hearer mistakenly attributing a message to the speaker[67] when the speaker never

---

[66] My focus here will be on our understanding of the content of utterances. However, our understanding of the types of act performed can also be shaped by misleading stereotypes. For example, we may mistakenly interpret an order as a request. Such cases are discussed in Kukla (forthcoming) who argues that the very acts we perform are determined by their actual uptake, meaning that certain groups are limited in their ability to perform certain acts (such as issuing orders).

[67] Or making a similar mistaken content related judgement, since it is not clear that audiences will always attribute one

intended to convey that message. That is, interpretative injustice occurs when prejudicial stereotypes result in miscommunication. As we will see, these prejudicial stereotypes affect both speech perception and the assignment of meaning. Thus, interpretative injustice is a broad phenomenon. I mostly follow Fricker's use of the term 'prejudicial stereotype' to mean, roughly, a generalisation which embodies a judgement about a social group, which is not properly evidence responsive[68]. That is, one harbours a prejudicial stereotype if one harbours a (usually negative) generalisation (or a set of associations which embody a generalisation) about a particular social group, and one's harbouring of that generalisation (or set of associations) is not sensitive to the available evidence. This constitutes an injustice for the following reason: If an aspect of our social practice systematically, unfairly, and disproportionately harms certain social groups, then in cases where members of those groups have been harmed by the social practice they have been victims of an injustice. Interpretative injustice gives rise to numerous harms (discussed in detail in section three), and these harms are experienced disproportionately by particular disadvantaged social groups. Moreover, in cases where these harms are experienced as a result of interpretative injustice it will be the hearer's lack of evidence responsiveness which gives rise to the harm. Thus it seems that the harm is unfairly inflicted upon the speaker (the miscommunication does not result from either a fault in the speaker or mere bad luck, rather it arises out of a fault in the hearer, or the wider social situation in which the hearer is embedded)[69]. Moreover, these harms will often constitute part of a wider web of harms which the disadvantaged individual is systematically subjected to. On some views of assertion (for example, Cappelen and Lepore (2004)) a typical assertion will have a huge number of propositional contents, many of which bear no relation to the speaker's communicative intentions. Thus it is worth noting that interpretative injustice relates to the audience's recovery of the speaker's intended proposition (or a similar message) rather than an asserted content more generally. If Cappelen and Lepore's speech act pluralism is correct then recovery of an asserted content does not rule out interpretative injustice.

It is worth flagging some similar notions which I will put aside for the rest of the chapter. Firstly, I do not consider interpretative injustice, as discussed here, to include cases where a hearer reaches the correct judgement about what a speaker intended to communicate, but does so on the basis of a

---

particular message to the speaker.

[68] It was mentioned earlier that characterising stereotypes as generalisations may not be optimal, since the processes underlying both linguistic understanding and our credibility judgements are likely to be associative. However, we can think of sets of associations as embodying generalisations. For example, the association of 'black' with 'crime', 'gun', and 'drugs' would embody a generalisation about black people like 'black people are criminals'.

[69] This is not to say that audiences in cases of interpretative injustice are always completely free of responsibility for the miscommunication. An already unclear speaker who's chances of being understood are futher undermined by the audience's prejudices is still at an unfair communicative disadvantage as a result of the audience's prejudices.

misleading stereotype. Such cases seem to be instances of lucky interpretation, and may be interesting in their own right. However, I will not be discussing them further.

Some cases involving luck do fall under the banner of interpretative injustice though. For example, consider the following case in which the audience does not luckily reach the correct interpretation, but is lucky in the sense that their prejudice does not prevent them from understanding an utterance they otherwise would have understood[70]: A person of colour tries to subtly communicate *p*, and does so in such a way that it might be unclear to even a reasonable audience that *p* was intended. However, the audience simply considers persons of colour to be incapable of subtlety, and so straightforwardly interprets the speaker as asserting *q*. In this case although the chance of being understood was already reasonably low the audience's prejudices lower the chances even further, in a way which is out of the speaker's control. This is analogous to throwing away someone's lottery ticket on the basis of their race. It is unlikely that they would have won even if the ticket had not been thrown away, however their chances of winning are substantially and unfairly diminished further by the act of throwing away the ticket. It is not clear that the audience would have understood the utterance even if they had not been prejudiced, however this is still a case of interpretative injustice.

Also of potential interest is the more general case where someone misinterprets a communicative act on the basis of a false belief or presupposition. Interpretative injustice is a subspecies of this phenomenon, and many of the things I say about the harms of interpretative injustice will carry over. However, I am concerned primarily with interpretative injustice on the basis of prejudicial stereotypes or associations because it is a more systematic phenomenon.

Finally, it is worth distinguishing interpretative injustice from Fricker's notion of hermeneutical injustice. Hermeneutical injustice is the phenomenon whereby a subjugated group or individual is unable to render their experience intelligible either to themselves or to others, due to their lacking the concepts with which to communicate (or render intelligible) their experience. Fricker uses the concept of sexual harassment as an example. Until the notion of sexual harassment entered into public discourse women who experienced sexual harassment were severely limited in their ability to communicate the species of wrong they were being subjected to, or even to identify it themselves in a clear way (Fricker (2007)). Interpretative injustice does not concern the existence or availability of the concepts required to render our social experiences intelligible. Rather, interpretative injustice

---

[70] Thanks to an anonymous reviewer for providing this case.

occurs when the wrong content is assigned as a result of prejudicial stereotypes influencing interpretation, regardless of what concepts are available in our public language. The existence of interpretative injustice is therefore entirely consistent with the absence of hermeneutical injustice. With that said I will turn to the task of arguing that interpretative injustice is a genuine, and genuinely problematic phenomenon.

In cases of context sensitivity, loose talk, and implicature, there are metaphysical and causal/epistemic questions about what is said or implied. That is, there is the question of what it is for a particular proposition to be implied, or for a particular proposition to be asserted (the metaphysical question), and there is the question of how it is that we, as audiences, come to judge that a particular proposition has been asserted or implied (the causal/epistemic question). The causal/epistemic question will be more important for us here, since we are concerned with the propositions attributed to speakers by audiences. However, the metaphysical question is still important. This is because if the facts which determine whether a particular proposition has been asserted include facts about the speaker and hearer's representations of one another's cognitive states, then factors which can have a harmful effect on these representations will have a knock on effect on the metaphysics of what is said.

As noted in previous chapters, we judge a context sensitive term to receive a particular value in context, a particular proposition to be implied, or loose talk to be resolved in a particular way, by making judgements about the conversational common ground and mutual salience. That is, we make judgements about what information the speaker trying to communicate, and the information to which they expect us to appeal, on the basis of our representation of the speaker's representation of our mutual goals, knowledge, interests, intelligence, salience judgements etc. This follows from a very simple view of the aims of communication: when we communicate we generally aim convey our intended meaning to the audience, and the audience generally aims to recover the speaker's intended meaning. Given these aims it makes sense that both audience and speaker would plan their utterance, or interpret on the basis of their best judgements regarding the other interlocutor's representation of the factors which would lead to a particular proposition being interpreted as asserted or implied. The same line of reasoning supports the notion that we rely on information about the common ground when we are confronted by unfamiliar accents, or noisy environments. In such situations we have to do some work to calculate the speaker's meaning, so we will appeal to hypotheses about what the speaker likely intended to communicate.

Indeed, some philosophers think that conversational common ground plays a metaphysical role in the determination of what is said as well. For example, Jeff King (2013, Forthcoming A, Forthcoming B) has argued that the value a context sensitive term receives in a context is the value which A) the speaker intends, and B) a competent interpreter who knows the conversational common ground would attribute. Whether a competent interpreter who knows the common ground would attribute a particular value to a term depends on what propositions are included in the common ground (that is, the set of propositions mutually known to be presupposed by speaker and audience).

However, it is rare that we explicitly reason about the common ground, or about mutual salience. Rather, like our credibility judgements our judgements about what is said or implied are unreflective and intuitive. We don't explicitly consider the common ground and work out the speaker's meaning in an effortful step by step manner, rather we rely on quick heuristics, and associations or stereotypes in order to quickly make a judgement about the common ground and about the asserted or implied proposition. As in the case of credibility judgements a hearer's stereotypes concerning social identity will have an impact on the judgements they reach. For example, a stereotype concerning social identity could alter the audience's representations of the speaker's likely interests or background knowledge.

Indeed, there is even empirical support for the notion that social identity judgements play a role in semantic processing. For example, Van Berkum et al (2008) found that anomalies with respect to the social identity of the speaker (for example 'I have a large *tattoo* on my back', spoken in an upper class accent, or 'If only I looked like *Britney Spears* in her latest video' spoken in a male voice (Van Berkum et al, (2008), p 2)) yielded the same type of neural response as semantic anomalies (such as 'the earth revolves around the *trouble* in a year'). This suggests that speaker identity plays a role in semantic interpretation even at the earliest stages. More recently Boland and Clark (MS) found that in linguistic contexts which promote predictive processing audiences were able to more quickly assign a referent to an ambiguous term when the referent was associated with the social identity of the speaker (for example, finger nails were more quickly identified as the referent of 'nails' when the sentence was spoken in a female voice). Finally, Geiselman and Bellezza (1977) found that audiences judged the same sentence to be more or less 'potent' (a measure of connotative meaning) when spoken by a male or female respectively. These studies support the view that amongst the stereotypes upon which we rely to recover content are stereotypes concerning social identity (e.g. class, gender, race etc.).

It is worth considering a specific account of the way in which we assign contextual values in order to get clear on precisely how implicit stereotypes are able to play a role. I use relevance theory as an example because it is a clear and relatively well known account of the mechanisms underlying the resolution of context sensitivity. According to relevance theory we assess interpretations for relevance, and select the first interpretation to meet our expectation for optimal relevance. 'Relevance' is a term of art denoting a feature of inputs to cognitive processes which strikes the greatest balance between relevant cognitive effects (new derivable information, and strengthening or weakening of previous assumptions) and cognitive effort (Sperber and Wilson (1986), (2004), Carston (2002)). That is, when we assign a meaning to a term we rank concepts on a basis of cognitive activation and assess each for relevance. The first concept to meet the audience's expectation of relevance is assigned as the meaning of the term. Consider the sentence 'Betty walked to the edge of the cliff and jumped'. Imagine that it is used in a discussion concerning Betty's suicide. The audience may start by assessing the literal meaning of 'jumped', but the resultant proposition would not meet the audience's expectation of relevance, so they would move on an assess other candidate meanings. The topic of conversation (suicide) will make the notion of 'jumping of a cliff' salient, and will thus be amongst the potential meanings assessed for relevance. It would meet the audience's expectation of optimal relevance, and so would be assigned as the meaning of 'jumped'. There are multiple points at which stereotypes and generalisation based on the speaker's social identity can enter into and effect this process. For example, an audience's view of the speaker will effect the level of activation of the various concepts which are ranked in order to be assessed for relevance. That is, if an audience member associates a particular concept or topic with a particular race, gender, or social class then a candidate meaning related to that concept or topic may be assigned a higher initial ranking when the speaker is a member of that race, gender, or social class. Likewise, the cognitive effects taken to be relevant will be affected by the audience's expectations regarding the aims and interests, and beliefs of the speaker.

Indeed, such factors don't just guide us in our resolution of context sensitivity, loose talk, and implicature. By observing our informants and making judgements about their beliefs, backgrounds, and interests, we build up a picture of the speaker which also guides our expectations regarding what words they are likely to use. There is reason to think that social identity judgements play a role in speech perception - in our judgements about the very words spoken. I am sure every reader will have found themselves in a situation where they mistook one word for another on the basis of expectations about what the speaker was going to say. For example, picture yourself in a noisy bar

talking to a well groomed, well spoken man in an expensive looking suit. Suppose he is actually a social scientist (you don't know this), and he says 'I work on current opinions on markets'. Due to the noisy environment you could easily mishear him and, partly on the basis of the expectations you developed as a result of his appearance, hear him as saying 'I work on currency options markets'. If he were less well dressed and had a more working class accent you may mishear him as saying 'I work currently in an open market' (as in, an open air market). These are clear cases of social identity judgements leading to miscommunication (for psychological and linguistic research into the impact of social information (including social identity judgements) on sentence processing see (Casasanto (2008), Campbell-Kibler (2010), Creel and Bergman (2011), and Sumner et al (2014)))[71],[72].

Sub-personal reliance on stereotypes about the speaker is an essential feature of utterance interpretation. However, as we have already seen such reliance on stereotypes can lead us astray. One can imagine many ways in which prejudicial stereotypes in particular can lead to problems. For example, Payne (2001) observed that white subjects primed with images of black faces were more likely to misclassify an image of a pair of pliers as an image of a gun. You might imagine a similar situation arising with speech perception. For example, white subjects may be more likely to mistakenly hear a black speaker as saying 'I've got a gun', when saying 'I want some gum'. It is not hard to see how such misinterpretation, if common enough, could be very problematic (for example, in cases involving trigger happy police officers). One can also imagine cases where someone's testimony is not understood and, as a result of prejudicial stereotypes, the audience just assumes the testimony was irrelevant. For example, imagine a rather unintelligent and prejudiced secondary supply teacher covering a social studies class. A black male student invokes the notion of hegemony whilst making a point. However, the teacher is not aware of the concept of hegemony, she just assumes that the student is referring to some aspect of urban culture which has little bearing on the class. She thereby writes of his testimony. This would be a case of interpretative injustice since it pertains to the type of content attributed to the student rather than the teacher's assessment of the student's credibility. She may take him to be a perfectly credible informant with respect to urban culture (the topic to which she thinks he is referring), but is simply not interested in

---

[71] The phenomenon of stereotypes and generalisations infecting our speech perception, can be seen as analogous to the phenomenon of cognitive penetration in the philosophy of perception. Cognitive penetration occurs when background beliefs, expectations, and emotional states affect the character of our visual experiences. For example, if we expect someone to be angry we may see them as angry, or if one has an overwhelming fear of guns and glances a long dark object out of the corner of one's eye one may perceive it as having gun like features (see Siegel (2012, 2013) for discussion of the epistemic significance of such cognitive penetration).

[72] Just to be clear, I don't intend this to be taken as a case of interpretative injustice, for no harm is brought about and the biases leading to the misinterpretation are not prejudicial.

information about urban culture.

Clearly the problems just outlined are important - the first due to it's ability to disproportionately subject certain groups to potentially harmful situations, and the second because it constitutes a form of silencing. However, I would like to focus primarily on a more general problem raised by prejudicial stereotypes in interpretation. The problem is simply that, as a result of systematic misleading stereotypes, utterances by members of certain groups are going to be misinterpreted far more often than the utterances of other groups. Accurate stereotypes (or, stereotypes which serve as reliable heuristics) will generally aid hearers in recovering what the speaker intends to communicate, in the same way that reasoning based on accurate generalisations will usually lead us to the truth. However, if there are popular generalisations, stereotypes, or associations about certain groups which get things radically wrong then the application of these stereotypes in utterance interpretation will lead to these groups being disproportionately misinterpreted. This is simply due to the fact that reasoning on the basis of faulty generalisations will generally lead to faulty beliefs. This in and of itself gives rise to certain harms[73].

The problems don't end here however. In her discussion of credibility judgements Fricker notes that stereotypes of disadvantaged groups have often included negative generalisations about intelligence, rationality, and knowledgeableness. Disadvantaged groups are often perceived as unintelligent, irrational, and lacking in knowledge. However, judgements about the likely knowledge, rationality, and intelligence of the speaker will determine the charitability of the audience's interpretation. That is, if one's interpretation is based on the assumption that the audience is unintelligent and has various false beliefs then one's interpretation will be less charitable than if one interpreted on the basis of the assumption that the speaker is a rational and reliable belief former. The result is that members of disadvantaged groups are not only more likely to be misinterpreted by members of more powerful groups, but are more likely to be interpreted The fact that certain groups are disproportionately subject to uncharitable interpretation is evidenced by a recent study by Alison Brooks et al (Forthcoming). Brooks et al found that potential investors systematically preferred entrepreneurial pitches by men over identical pitches by women, and that

---

[73] Indeed, if King (2013, Forthcoming A, Forthcoming B) is correct then these problematic stereotypes might lead to nothing being said at all in certain cases. If one's prejudicial stereotypes lead to false judgements about what the speaker does or does not take to be common ground, then these prejudicial stereotypes will influence which propositions are in the common ground. Therefore, they could result in the information the speaker intends the audience to appeal to being absent from the common ground. The result of this is that a competent interpreter who knew the actual common ground might not be able to recover the speaker's intended meaning. Thus, on King's account, no proposition would be asserted.

they preferred pitches by attractive men to identical pitches by unattractive men. Participants were asked to rate how persuasive, fact based, and logical the presentations were, and found presentations by men (especially attractive men) to rate higher on each scale, even though the female and male entrepreneurs were reading from the same script. Since the extent to which a presentation is logical or fact based is dependent on the content of the presentation it seems reasonable to conclude that the gender of the entrepreneur had an impact on the way they were interpreted, with males being interpreted more charitably, and attractive males having a particular advantage. Charitability will be especially relevant when we reconsider Fricker's response to epistemic injustice. First, however, it is worth spending some time considering a case in which these problems arise, and clarifying precisely why the problems just outlined give rise to harms.

3. *The Harms of Interpretative Injustice.*

In section two it was argued that interpretative injustice constitutes a wrong because it systematically and unfairly harms members of disadvantaged social groups. We have seen how the phenomenon might arise. However, we have not yet considered any detailed examples, nor have we considered any specific harms which will be systematically associated with interpretative injustice. In this section I will provide a clear example of interpretative injustice and to briefly highlight some of the ways in which interpretative injustice leads to genuine harms[74].

One of Fricker's central cases is the trial of Tom Robinson in Harper Lee's 'To Kill a Mocking Bird'. Tom Robinson is a black man falsely accused of rape in the deep south of the United States in the mid 30s. In Lee's tale Tom Robinson is wrongly convicted of rape despite the fact that a convincing case for his innocence has been put forward. Fricker uses this as an illustrative example in which an audience (the jury) do not assign an accurate degree of credibility to the speaker as a result of their prejudicial biases, and the speaker is harmed as a result. As Fricker notes, in the racial climate in which the trial takes place it is very difficult for the jury to take the word of a black man over the word of a white woman, especially in a rape case. This was, after all, a period in which there was still a very widespread and open anxiety about the sanctity of white southern womanhood and the supposed threat of predatory, animalistic black male sexuality. Considering the trial of Tom

---

[74] I am not concerned with how 'epistemic' each of these harms is. Certainly they are epistemic in the sense that they contribute to Fricker's form of epistemic injustice - they relate to our ability to share knowledge, and they involve people being wrongly held epistemically responsible for certain propositions. However, it is less clear that the speaker is directly wronged as a knower in instances of interpretative injustice. Indeed, in many of these cases the problem arises because the speaker is treated as a knower, but treated as a knower of the wrong proposition.

Robinson, and its historical context, we can come up with clear cases of interpretative injustice. Fricker draws our attention to a particular passage in which Robinson is asked why he visited his accuser's home regularly. Robinson did so because he felt sorry for his accuser, she seemed to live a lonely sad life, and he wanted too help her. He expresses this by saying that he felt sorry for her. This goes down very badly. In the climate of the trial the very notion that a black man could feel sorry for a white woman (who would be considered his superior) would be found shocking, and the audience would likely read into his utterance that he felt superior to her. This, in and of itself, might be seen as a case of interpretative injustice. However, I am primarily concerned with cases where the primary communicated content of an utterance is misinterpreted, and I think we can modify this case to get a clearer example of interpretative injustice along such lines. Let us suppose that Robinson did not state that he felt sorry for his accuser, but rather that he visited her regularly because she seemed vulnerable. Here the term 'vulnerable' could be read different ways. As used by Robinson it would have meant (or been intended to mean) 'in need of help/assistance'. However, in the context of the racist south where one of the dominant associations of black maleness was an association with sexual aggression (especially toward white women), as well as being uncontrolled, animalistic, predatory, uncivilised, and morally inferior, the description of the apparent victim of a rape as 'vulnerable' could easily be seen as him accidentally revealing that he saw here as an easy target for sexual violence. Indeed, if he realised his misstep here and 'shifted uncomfortably in his chair' (as he does in the book), this could be taken as further confirmation that he has just accidentally reviled his true intentions in the case – slipped up by revealing that he saw her as an easy target[75]. I am sure it is clear to the reader that misinterpretations such as this can lead to great harm, and that a group's being systematically and disproportionately subject to such misinterpretations constitutes a social injustice. For the rest of this section I will outline three particular harms which seem to arise from interpretative injustice (and which, thus, are inflicted disproportionately and unfairly on certain groups).

Now that we have considered a clear case of interpretative injustice it is worth considering several harms which systematically arise from interpretative injustice. Firstly, interpretative injustice makes it harder for some groups (particularly disadvantaged groups) to be heard. That is, certain groups are silenced disproportionately. Silencing occurs when an individual or group is prevented from carrying out a communicative action, either by being prevented from attempting the action (for example, through intimidation) or as a result of the action being prevented from being successful.

---

[75] One can see how this would fit into the relevance theoretic story given earlier. The relevant associations of predatory animalistic sexuality will be highly salient to the audience, thus the resultant concept of vulnerability will receive a higher level of activation and thus be ranked above Robinson's intended meaning.

Hornsby and Langton (1998) characterise the latter forms of silencing in terms of the illocutionary or perlocutionary force of the act being blocked. That is, either the communicative act is prevented from producing its intended outcome (perlocutionary silencing) or it is prevented from even constituting the intended type of action (illocutionary silencing). Hornsby and Langton maintain that in order for an illocutionary act (for example, an assertion or a refusal) to be carried out the fact that it is intended to be an assertion or refusal must be recognised by the audience. So when the audience fails to recognise the illocutionary act the speaker is attempting the speaker is illocutionarily silenced[76]. More recently Ishani Maitra[77] (2009) has characterised silencing in Gricean terms. Grice's conditions on speaker meaning are as follows:

'A speaker *S* means something by uttering *x* iff, for some audience *A*, *S* utters *x* intending:

(i) A to produce a response *r*;

(ii) A to think (recognize) that *S* intends (i); and,

(iii) A's fulfilment of (ii) to give him a reason to fulfil (i)'

Grice (1989), 92.

Maitra then characterises silencing as follows:

'On my view, a speaker is communicatively disabled iff she is unable to fully successfully perform her intended communicative act, because her intended audience fails to satisfy either the second or the third of her (Gricean) intentions.' Maitra (2009), 327-328.

All cases of interpretative injustice are cases of silencing, since if one misinterprets the content of an utterance then one thereby fails to recognise the illocutionary act being performed, or the speaker's communicative intention[78]. Indeed, the paradigmatic case of ethically problematic silencing[79] is arguably a case of interpretative injustice. In the paradigmatic case of silencing a

---

[76] As mentioned in footnote 6 Jeff King's (2013, forthcoming a, forthcoming b) metasemantics entails a similar form of silencing - if the common ground is rendered defective in such a way that the information contained therein would not be sufficient for a competent audience to recover the intended proposition then the speaker will be prevented from performing the act of asserting their intended proposition.

[77] Maitra is not aiming to capture cases of silencing where the speaker is prevented from even attempting a communicative act.

[78] It is not clear that all cases of silencing in Hornsby and Langton's sense will be cases of interpretative injustice, since one could misinterpret the illocutionary force of an utterance without misinterpreting the content. However, these cases will still be related, since they will often be cases in which some other aspect of an agent's speech act (the actual act being performed) is misinterpreted, perhaps on the basis of problematic stereotypes.

[79] Maitra (2009) argues that all silencing is ethically problematic. I find this claim problematic. It is not clear that the speaker is wronged when it is their own fault that their communicative intention has not been recognised.

woman tries to refuse a man's sexual advances by saying 'no'. However, as a result of his misleading conception of women as always wanting to avoid the appearance of promiscuity, yet really usually desiring sex, he misinterprets the woman's refusal. He thereby fails to recognise her communicative intention. Thus the woman is prevented from performing the illocutionary act of refusal. In this case the man's interpretation of the woman was guided by a misrepresentation of the speaker's interests, desires, and intentions based on a harmful stereotype of women. As emphasised by Maitra (2009), the harms of silencing do not end here. She writes:

> 'Speech has, at the very least, great instrumental value. It enables us to get what we want and need, for ourselves and others. It constitutes our first line of defence against a variety of injuries, from unwanted sexual overtures to tyrannical governmental action. And it is essential both to the propagation of knowledge, and to the proper functioning of a democratic society. When a speaker is communicatively disabled, she is thereby deprived of these (and other) benefits that speech can offer.' Maitra (2009), 331.

As Maitra points out, when one is silenced one is thereby restricted in one's ability to share knowledge. Fricker, in her discussion of epistemic injustice, emphasises the importance of the ability to convey knowledge. Firstly she argues that the ability to share knowledge is a fundamental human value, and that when a group is limited in their ability to share knowledge they are thereby limited in their ability to engage in a practice fundamental to human value, and to human society more generally. If prejudicial stereotypes lead to some groups being misinterpreted (and therefore silenced) disproportionately then the prejudicial stereotypes make it harder for these groups to enter into the practice of sharing knowledge. Thus, if Fricker is right, these stereotypes will make it harder for certain groups to engage in social practices fundamental to human value. Moreover, it will make it harder for these groups to pursue their ends[80]. For example, it will make it harder for individuals in certain groups to progress in their careers. Brooks et al (forthcoming) illustrated one particular way in which this could occur - female entrepreneurs are at a distinct career disadvantage, seemingly as a result of their tendency to be interpreted uncharitably (at least in certain situations).

Additionally, communication is a messy affair, and communicative success is always a matter of degree. I think it is likely we rarely have a 100% accurate grasp of the speaker's communicative intention. However I think it is implausible that we commit the wrong of silencing in the majority of communicative interactions.

[80] A related point is that as a result of stereotypes about a group the group itself (not just the individuals in the group) may be misinterpreted in political discourse. For example, if the poor are seen as generally being lazy and dependent on handouts, then the desire of the poor for greater welfare may be interpreted in political discourse as a desire for handouts. Thus, stereotypes about the group make it harder for the group's interests to be represented in a democratic system (thanks to Sebastian Becker for this point). A full account of this phenomenon would require an account of what it is for a group as a whole to have particular interests, and to communicate them successfully.

It is easy to imagine how the same phenomenon could occur in job interviews, marketing pitches, academic research presentations etc. In such cases the female speaker will be prevented from communicating her intended message, and will be assessed and treated as if she has attempted to communicate something different (in a similar way to the black male who is not understood when he tries to invoke the notion of hegemony). The case of Tom Robinson provides a particularly vivid example of harmful silencing. As a result of his inability to be understood he is wrongly convicted of rape, and (the book implies) is murdered soon after.

Since all cases of interpretative injustice seem to be cases of silencing it may be unclear why interpretative injustice is a worthy topic of investigation in its own right[81]. Why not just study the general phenomena of silencing? I think there are several reasons why it is important to study interpretative injustice as an independent phenomenon. Firstly, as we will soon see, there are distinctive additional harms not systematically caused by other forms of silencing[82]. Two of these harms will be discussed later in this section, and it seems plausible that there may be further harms which are caused only by interpretative injustice. Secondly, interpretative injustice complicates our response to epistemic injustice. This is discussed in the final section of the chapter where I argue that the phenomenon of interpretative injustice may render Fricker's solution to epistemic injustice suboptimal. The final (and more general) point is that although silencing is a unified phenomena, different types of silencing have different types of causes. In order to be able to find solutions to the problems caused by silencing, and in order to properly understand the way in which different social practices silence, it is important to have a clear taxonomy of the types of silencing not just in terms of their effects (i.e. whether they are perlocutionary, illocutionary etc), but in terms of their causes. The phenomenon of interpretative injustice fits into such a taxonomy because it has a distinctive cause – underlying prejudicial biases shaping the way we interpret people's speech. Other types of silencing have different causes, for example the same biases altering our credibility assessments, or more overt prejudices causing us to discount, or try to actively prevent the testimony of certain groups. Some types of silencing are unified by, for example, being caused by prejudicial stereotypes. However, these stereotypes can act in different ways on different levels of cognition, giving rise to importantly different forms of silencing. A solution to one of type of silencing will not necessarily generalise to all.

---

[81] Thanks to an anonymous reviewer for raising this concern.
[82] This is not to say that the harms to be discussed shortly are never caused by other forms of silencing. Rather, interpretative injustice, by its very nature (and unlike silencing in general), seems particulaly strongly (and systematically) associated with these particular harms.

Related harms occur when some groups are disproportionately interpreted overly charitably. In these cases it is not clear individuals are directly harmed. However, a practice of systematically interpreting some groups overly charitably (and interpreting other speakers with a more reasonable level of charity) will lead to some groups having specific advantages over other groups in a way that produces many of the same harms as silencing. For example, silencing makes it harder for certain groups to advance in their career. Overly charitable interpretation of one group will likewise make it harder for members of other groups to advance in their careers, because it will be harder from them to compete. A possible example of this occurring in philosophy might be philosophers from prestigious institutions being interpreted more charitably than philosophers from lower ranked institutions. It is easy to see how this could occur. Given the difficulty of strictly interpreting even the clearest analytic philosophy, there are usually a few ways an argument can be read. If our knowledge of the prestige of the author influences our interpretation of their argument this will place that author at an advantage over others who we may interpret less charitably (and perhaps more accurately) as a result of the status of their institution (race and gender could conceivably lead to the same form of overly charitable interpretation)[83],[84].

Secondly, when a speaker is interpreted as having asserted a proposition *p* they are held publicly accountable for defending *p*. This facts is emphasised by Macfarlane (2005, 2011), Hinchman (2005), Goldberg (2006) Rescorla (2009), and McMyler (2013). Hinchman puts the point as follows:

> "If you've told someone that p, that person is now under certain conditions -for example, if he's challenged whether p-entitled to hold you accountable for producing a reason to believe that p.' Hinchman (2005), 568.

Of course, in cases of misinterpretation (where the fault lies with the audience rather than the speaker, as is the case with interpretative injustice) the speaker does not actually say what they are interpreted to have said. So the audience will not actually be entitled to hold the speaker accountable for having said *p*. Nonetheless, if the audience thinks the speaker has said that *p* they

---

[83]   Of course, judgements based on institutional affiliation may not be completely unreliable, since philosophical ability has at least some role in candidate job placement. I do not wish to take a stand on the usefulness of institutional affiliation as a guide to ability here.

[84] Tom Robinson's trial illustrates another sense in which someone may be interpreted overly charitably to the detriment of a member of a disadvantaged group. If Robinson's accuser slipped up and accidentally implied that she had sexual desires toward Robinson, yet the jury were unable to interpret her (a southern white female) as expressing desire for a black male, then this would surely harm Robinson's chances of receiving a fair trial. And in this context this would constitute overly charitable interpretation (since the alternative interpretation would harm the accuser's case).

will hold the speaker responsible for *p*, even if they are not entitled to do so. If prejudicial stereotypes result in certain groups being misinterpreted more frequently than others (without the misinterpretation being the fault of the speaker) then members of these groups will also be disproportionately held responsible for intending to communicate things they never intended to communicate. Being held responsible for something you never did constitutes a harm.

Moreover, in combination with other prejudices this could lead to further problems down the line. In her discussion of credibility judgements Fricker notes that some groups have been perceived as being inherently dishonest. If this is the case then the disadvantaged speaker's ability to correct the audience's mistake may also be harmed. They could be perceived as being dishonest and trying to slip out of a commitment. Indeed, if some groups find themselves having to correct mistakes in interpretation more often than others this could add to the perception of that group as being sneaky and dishonest. This in turn will contribute to the stereotype that the group is dishonest. Thus, it will give rise to lowered credibility judgements.

Relatedly, victims of interpretative injustice will often be forced to defend themselves against misinterpretation (both during and following communicative exchanges). This experience of having questions raised against oneself, having one's credibility brought into question, and having to actively re-shape the audience's conception of ones-self, will often be cognitively and emotionally draining, and often humiliating. This will be especially true of cases in which one's words, which are one's primary means of shaping the audience's conception of oneself, are at risk of being misinterpreted and actually contributing to the conception one is trying to correct. We can imagine an extension of the Tom Robinson case in which this occurs. Imagine than Robinson is asked to clarify what he means by 'vulnerable'. Robinson may have replied 'well, she seemed like she needed a man'. It is unclear what should be built into the concept of 'man' here (presumably not any male will do), and it raises the question of the purpose for which a man is needed. In this case Robinson would have meant that his accuser needed the help of someone who was capable of performing manual labour around the house. However, in making this utterance he would once again be at risk of misinterpretation (due to the very same biases which lead to the misinterpretation of 'vulnerable'). The audience's conception of black men as animalistic sexual predators, and their preconceptions of Robinson's likely aims in the scenario, could lead to him being interpreted as stating that his accuser desired a dominant sexually aggressive black male (perhaps even with the implication that she desired the rape). This, of course, is far from what Robinson intended. However, after his two misinterpreted utterances any further attempts to clarify his assertions would

likely be taken as attempts to wriggle out of what he has said. I'm sure it is clear to the reader how humiliating, frustrating and emotionally draining this experience would be for Robinson (and not only because of the likely consequences of a rape conviction).

Finally, if members of certain groups are interpreted uncharitably more frequently than others then this will contribute to the stereotype that members those groups are poor informants, thereby feeding into the stereotypes which give rise to prejudicial credibility judgements. Consider the following: suppose we have two informants, Max and Sally. If we frequently interpret Sally uncharitably so that, for example, we take her to be asserting that *p*, where *p* is less plausible than some alternative interpretation *q* which we would assign to Max in the same circumstances, then the following situations will arise more frequently with respect to Sally than with respect to Max: A) we find Sally's utterance so initially implausible that we lower our judgement of her credibility[85][86], and B) we believe her only to later discover that the proposition we took the her to be asserting was false, which once again leads us to lower our judgement of her credibility. Therefore, as a result of our uncharitable interpretations of Sally we will find ourselves with what seem like good reasons for assigning Max a higher credibility than Sally, even though he may be no better as an informant.

If audiences make credibility judgements on the basis of the speaker's social identity, and the speaker's social identity is a partial determiner of the charitability of the audience's interpretation, then that audience will find themselves in situations A and B more frequently with respect to the testimony of certain social groups than others. This will thereby strengthen the stereotype that particular groups have low credibility, and thus contribute to prejudicial credibility judgements. We can imagine this occurring in the context of entrepreneurial pitches as discussed above. If an investor frequently has the impression that female entrepreneurs present weaker, less coherent and more illogical pitches, then they may come to the belief that female entrepreneurs are generally not very intelligent or reliable. As a result they may assign a lower credibility to certain assertions made by the entrepreneur, for example assertions which concern the entrepreneur's ability to carry out the project, or the entrepreneur's statements regarding projected profits. Thus, interpretative

---

[85]   Karen Jones (2002) discusses the potential harms produced when we lower our credibility judgement of the speaker on the basis of what we take to be the initial plausibility of their assertion.

[86]   It might be thought that we would not assign initially implausible interpretations if we are taking the gricean maxims to be in effect. This would be a mistake. When applying gricean norms you need to represent what would be a cooperative contribution given the speaker's representation of the situation. For example, imagine you know that p, and believe that a speaker doesn't know that *p.* If the speaker were to make an assertion which would imply *q* only if *p* was common knowledge, then you would not take them to be implying *q*. If you take the speaker's representation of the communicative situation to be defective then the contribution you take the speaker to be making will not be the most cooperative given the actual facts, but rather the most cooperative given a particular set of faulty background assumptions. And this contribution may seem implausible.

injustice contributes to epistemic injustice.

So far everything I have said can be seen as supplementary to Fricker's account of testimonial injustice. In the final section I argue that the two phenomena interact in an important way. More precisely, I argue that adopting Fricker's proposed solution to testimonial injustice without also treating interpretative injustice will cause the hearer to be epistemically harmed, with only minimal benefits to the speaker.

4. Interpretative Injustice and Credibility Adjustments.

Fricker does not merely identify and describe the phenomenon of epistemic injustice. She also offers a strategy for overcoming it. She argues that the appropriate response to epistemic injustice is to develop the virtue of testimonial justice, whereby we gain a sensitivity to the sorts of situations in which our credibility judgements may be biased, and re-consider our judgements in such circumstances. In such situations we should not rely on our quick intuitive judgements, but should try to make an accurate and unprejudiced judgement. The hope is that we eventually become more reliable habitual judges of credibility, and that prejudicial stereotypes will no affect our credibility judgements. At this point we will display an instinctive sensitivity to actual signs that the speaker has low credibility, and not judge speakers to have low credibility on the basis of factors such as race and gender.

It might be thought that developing the virtue of testimonial justice would make us more reliable belief formers. For example, suppose that on reflection I realise that I have an underlying bias against a particular race, and that I assign members of that race a lower credibility than I should. Next time I have an interaction with a member of that race I actively reconsider my intuitive credibility judgement, and compensate upwards. The effect of this is that the speaker is no longer prevented from communicating their knowledge to me, and I no longer miss out on knowledge I can gain from this speaker.

This solution becomes less straightforward once interpretative injustice is taken into account. Consider the case of Max and Sally from the previous section. However, imagine that in this case we have not only been interpreting Sally uncharitably, but have also been assigning her a low credibility, and thus not trusting her testimony. After reading Fricker's 'Epistemic Injustice' and

reflecting on the way we assign credibility judgements we realise that we have been assigning Sally a low credibility on the basis of prejudicial stereotypes. As a result we no longer trust our intuitive credibility judgements of Sally. Instead we actively reason about her credibility. Sometimes this results in our rejecting her testimony for reasons other than prejudicial bias. However, on other occasions it results in us compensating upwards and assigning her a high degree of credibility. That is, in many cases when Sally makes an utterance and we take her to be asserting a proposition p, we go against our instinctive (but prejudicial) credibility judgement and assign her a higher credibility, thereby leading us to believe p. Certainly we will find ourselves believing what we take Sally to have said more often than we would have done otherwise.

It should be clear why this will lead to problems if we do not also adjust for interpretative injustice. Consciously assigning Sally a higher credibility will not necessarily affect the charitability of our interpretation[87]. The practice of assigning high credibility to uncharitably interpreted utterances will lead to unreliable belief formation, and will thus be harmful to the hearer. Moreover, in the cases where Sally is misinterpreted she will still be blocked from sharing her knowledge. The audience will bear an epistemic cost which is of little benefit to the speaker. This result applies more generally. If we treat epistemic injustice in the way Fricker suggests, without also treating interpretative injustice, then in many cases we run the risk of not only continuing to harm the speaker, but also harming ourselves.

To be clear, I am not claiming that Fricker's solution fails to treat testimonial injustice. If we raise our credibility assignments then we will avoid disrespecting the speaker as a knower. Thus one of the main harms of testimonial injustice will be removed. Rather, my point is cautionary. Treating testimonial injustice in the way Fricker suggests without also treating interpretative injustice will, in cases where both interpretative and testimonial injustice are present, put the hearer at epistemic risk, and only make a minimal positive difference to the speaker.

It might be thought that this is unproblematic, since the virtue of testimonial justice can be applied earlier on in the process of testimonial belief formation[88]. This would involve shifting gear from being a passive spontaneous interpreter to being an active interpreter. The first point to note here is that the problem raised above was only intended to illustrate the risks of treating testimonial

---

[87] It is perhaps a matter of contingent empirical fact the practice of actively adjusting one's credibility judgements will eventually change one's overall perspective of the speaker, which will also solve the problem of interpretative injustice. However it is an empirical question (to which we don't have an answer) whether this will be the case.

[88] Thanks to an anonymous reviewer for this observation.

injustice without also treating interpretative injustice. If one shifted gear earlier on then one would not be ignoring interpretative injustice. Thus, there is no tension with the point above. However, there are two further points which indicate that it may be less straightforward to shift gears with respect to one's interpretation than it is to do so with respect to one's credibility judgements. The first (minor) point is simply that although the processes underlying both interpretation and credibility judgements are subpersonal, credibility judgements seem more easily accessible to consciousness. We are more used to actively reasoning about a speaker's credibility during conversational exchanges than we are to reasoning about how to interpret them. The second point is that credibility judgements seem to be scalar in a way in which interpretation is not. That is, it might be thought that our credibility judgements fall somewhere on a scale, and in order to adjust our judgements we must merely shift where we locate the speaker's utterance on that scale. With interpretation there are multiple possible meanings which need to be ranked for plausibility given our knowledge of the context. Thus, the task of adjusting interpretation cannot be reduced to a simple heuristic such as 'shift credibility upward'. This is not to say that such adjustment is impossible. However, given the greater demandingness it is likely to have a greater detrimental effect on the conversational exchange than a credibility adjustment. As a result, a practice of adjusting this way will still leave certain groups at a communicative disadvantage compared to those for whom no adjustment is needed.

5. Conclusion.

I have introduced the notion of interpretative injustice and explained the numerous harms to which it gives rise. It gives rise to silencing, unjust attributions of responsibility, and it contributes to the stereotypes which give rise to epistemic injustice in credibility judgements. Moreover, I have argued that interpretative injustice causes a problem for Fricker. That is, if the virtue of testimonial justice is developed without interpretative injustice being treated as well then in many cases harm will be caused to the audience without any benefit to the speaker.

It is not clear how we should respond to interpretative injustice. The practice of actively reflecting on our interpretations whilst communicating would be costly. After all, we have developed fast sub personal mechanisms of interpretation for good reason. Indeed, if it is generally more cognitively demanding for members of certain privileged groups to communicate with those from less privileged groups then the less privileged individuals will still have problems integrating. I suspect

that the problems of interpretative and epistemic injustice cannot be treated alone, and are most efficiently treated by eliminating the harmful prejudicial stereotypes from one's psyche altogether. It is an empirical question how this is best achieved (if it is possible at all). However, one thing we can do in the mean time is to develop a sensitivity to the kinds of situations in which interpretative injustice may be especially harmful. If we develop such a sensitivity, and actively reflect on our interpretation in such situations, then at least some of the more severe harms of interpretative injustice will be stemmed.

Chapter Six: Cognitive Penetrability and Testimonial Anti-Reductionism.

0. Introduction.

This chapter explores testimonial versions of the cognitive penetrability cases presented by Susanna Siegel (2012, 2013). It is argued that the phenomenon of cognitive penetration (or an analogous phenomenon) can arise at the level of linguistic understanding (both as a result of speech perception, and interpretation). It is argued that a range of views about testimony (especially certain brands of anti-reductionism which purport to sustain a priori testimonial knowledge) make the wrong predictions about the cases. I start by outlining Siegel's cognitive penetrability arguments against perceptual dogmatism. Next I distinguish several forms of testimonial reductionism and anti-reductionism in order to indicate precisely which theories of testimony will be rendered problematic. Following this I outline cases of cognitively penetrated linguistic understanding arising out of both speech perception and linguistic interpretation, and I explain why these cases render certain views of testimony problematic. Finally I discuss several responses. I start by arguing that several responses to Siegel's arguments regarding perception do not carry over to the case of testimony. I then discuss three further responses specific to the case of testimony (one of which is found to have promise, but places difficult constraints on one's account of the basis for one's right to trust one's own linguistic understanding).

1. Perceptual Dogmatism and Cognitive Penetrability.

Perceptual dogmatism holds that when it perceptually seems to us that p we are by default justified in forming a belief that p. That is, perceptual seemings themselves provide a (defeasible) justification without need for any further knowledge or inference. Dogmatism has recently been challenged by Susana Siegel (2012, 2013). Siegel explores the hypothesis that our perceptual experiences can be cognitively penetrated. The general idea is that other doxastic and attitudinal states can have an impact on the contents of our perceptual seemings. That is, a state S_1 (such as a

perceptual seeming, or an intuition that *p*) is cognitively penetrated by a state *S₂* (such as a belief or a desire) if the content of *S₁* is caused (in some internal mental way) by S_2[89]. For example, Payne (2001) found that subjects under a time constraint were more likely to misidentify an image of a tool as an image of a gun when primed with images of black faces than when primed with images of white faces. There are several possible explanations for this. One such possibility is that the underlying expectations and associations evoked by the images of the black faces actually effected the subjects' perceptions of the tool, causing them to see it as a gun.

Perceptual dogmatism instructs us to form beliefs even on the basis of cognitively penetrated perceptual experiences, and this, according to Siegel, entails that some seemingly irrational beliefs will be justified. In Siegel's cases an irrational belief or attitude B_1 influences the content of the perceptual seeming, meaning that the resulting belief B_2 ultimately arises partly as a result of that irrational belief or attitude B_1. Siegel argues that we should treat the resultant belief as unjustified. It will help to consider one of Siegel's examples:

> ANGRY-LOOKING JACK: Jill believes, without justification, that Jack is angry at her. The epistemically appropriate attitude for Jill to take toward the proposition that Jack is angry at her is suspension of belief. But her attitude is epistemically inappropriate. When she sees Jack, her belief makes him look angry to her. If she didn't believe this, her experience wouldn't represent him as angry. Siegel (2012), 209.

Since it perceptually seems to Jill that Jack is angry she is, according to the dogmatist, justified in believing that Jack is angry. Experiences with such problematic causal origins are labelled 'checkered experiences'. Siegel hypothesises that such experiences do not provide justification. This is characterised by the following principle:

> DOXASTIC DOWNGRADE THESIS: If S forms a first-order belief B with content P, on the basis of an experience E that is checkered with respect to its content P, B is thereby doxastically unjustified, assuming that S has no other basis on which she believes P[90]. Siegel

---

89

  This definition is a rough generalisation of that given by Stokes (2012). The precise character of cognitive penetration is not of central importance to this chapter, as what really matters is just that background beliefs and attitudes can have an impact on the contents of states which are taken to provide default justification in such a way as to undermine that justification.

90 Siegel also provides a propositional version of the downgrade principle. However, she takes the doxastic

(2013), 704.

There are several reasons to endorse the doxastic downgrade thesis. Firstly, many find it counter intuitive to claim that beliefs (such as Jill's) which arise from checkered experiences are justified. Secondly, the features by which ill-founded beliefs transmit their ill-foundedness to other beliefs are, according to Siegel, shared with checkered experiences. Thus, there is reason to think that beliefs based on checkered experiences, just like beliefs based on other unjustified beliefs, are not justified. Siegel (2013) argues for this by going through several differences between beliefs and perceptual seemings, and arguing that none of these differences could account for the ability of beliefs, but not seemings, to transmit illfoundedness. If Siegel is correct then it is likely that the features by which ill-formed beliefs transmit their lack of justification are shared with checkered experiences. For example, beliefs differ from perceptual seemings in sometimes being adjustable in response to evidence, in being rational or irrational, and in being formed through explicit reasoning. However, according to Siegel, none of these features explain the ability of a belief to transmit irrationality to other beliefs. For example, implicit beliefs can transmit irrationality, as can beliefs which are very hard to adjust. Even rational beliefs can, according to Siegel, help transmit irrationality. Thus, it seems that whatever feature of belief allows it to transmit irrationality is shared with perceptual experience[91]. Finally, Siegel (2012) argues that cognitive penetration can produce viciously circular beliefs. Consider Jill's belief about Jack's anger again: Her conclusion that Jack is angry is dependent on her perception of him as angry, which is in turn dependent on her prior belief that he is angry. To illustrate this final point Siegel draws the following analogy:

> GOSSIP CIRCLE: In a gossip circle, Jill tells Jack that p, Jack believes her but quickly forgets that she's the source of his belief, then shortly afterward Jack tells Jill that p. It seems silly for Jill to take Jack's report that p as providing much if any additional support for p, beyond whatever evidence she already had. On the face of it, this looks like a feedback loop in which no new justification is introduced. Similarly, when beliefs are formed on the basis of cognitively penetrated experience, it is as if your belief that p told you to have an experience that p, and then your experience that p told you to believe that p. Siegel (2012), 202.

---

downgrade principle to be of more central importance.

[91] Of course, an argument like this can never be conclusive, as it is always open that a new feature of belief will be identified which explains its ability to transmit illfoundedness and which is not shared with states of perceptual seeming.

It seems plausible that these problems will generalise beyond perceptual dogmatism to other views on which cognitively penetrable states are taken to provide default justification. Perceptual seemings present the world as being a particular way, are taken by dogmatists to provide default justification, and appear to be cognitively penetrable. However, other sources such as memory and intuition also seem to present the world a particular way, can be taken to provide default justification, and may well be cognitively penetrable. If one's intuition that *p* is somehow caused by a prior belief that *p*, and is still taken to provide default justification, then this would give rise to a similar sort of problematic circularity (of course, whether memory or intuition actually are penetrable in this way is unclear).

2. The Varieties of Reductionism and Anti-Reductionism.

My primary aim will be to argue that cases similar to those given by Siegel cause problems for particular forms of testimonial anti-reductionism. However, strictly speaking these arguments can also be taken to apply to certain conceivable brands of reductionism. Furthermore, reductionist and anti-reductionist theories are sometimes formulated in such a way that they fail to make predictions about the cases I discuss. Thus, I will endeavour in this section to clarify several different ways of spelling out reductionist and anti-reductionist views of testimony, in order to get clear on which views are rendered problematic by cases of cognitive penetration.

The distinctions I draw concern different stances a theorist might take on the question of the conditions under which an agent has an epistemic right to trust in their own linguistic understanding. Thus, before getting into the distinctions it is worth saying a little about linguistic understanding. When we hear others speak we hear their words as words, not as mere sounds. We hear assertions not as mere noises but as assertions, with a particular content. It is as if we perceive utterances as having a particular meaning and force. Following Fricker (2003) let us call these apparent perceptions of meaning and force 'quasi-perceptions'. Quasi-perceptions appear to play a central role in the normal process by which we acquire testimonial knowledge. It is via our quasi-perceptions of meaning that we form beliefs about what is said. Arguably, veridical quasi-perception is necessary for understanding what has been said[92][93]. We will distinguish different

---

[92] It is unclear what constitutes a veridical quasi-perception of meaning, because it is unclear under what conditions a speaker counts as having said that *p*. Unfortunately there is not sufficient space to discuss the issue here.

[93] It is important to note that veridical quasi-perception is not required for knowledge of what is said, as you might find

views of testimony in terms of their different commitments regarding whether or not we have a default right to take our quasi-perceptions at face value.

Reductionism and anti-reductionism about testimonial justification are usually thought of as views concerning whether or not agents must, if their testimonial belief is to be justified, possess reasons to accept testimony. There are two importantly different ways in which such views might be formulated: They may be formulated as views about the conditions under which a testimonial belief is justified (that is, as theories about the epistemic status of the belief formed on the basis of testimony)(call these 'resultant belief' formulations), or they may be formulated as views regarding the conditions under which one is warranted in transitioning from the belief or knowledge (or fact) that a speaker has said that *p,* to a belief that *p.* We might characterise a transitional version of reductionism as the view that one requires reasons to trust the speaker in order to transition from 'S said that p' to 'p'. Anti-reductionism would be the denial that such reasons are required. Transitional views obviously have implications for the epistemic status of the resultant belief. If one needs reasons to make the transition, but one makes the transition without the requisite reasons, then the resultant belief will be unjustified. However, transitional views do not tell the whole story, as they say nothing about the agent's epistemic rights with regard to their own linguistic understanding. Transitional views will not be directly challenged by the arguments presented here, as the arguments I present concern whether or not one has a default entitlement to trust one's quasi-perceptions.

Reductionist and Anti-reductionist views are sometimes formulated not as views about justification but as views about knowledge. Anti-reductionism about testimonial knowledge would be the view that testimonial knowledge needn't be based on inference or positive reasons possessed by the speaker. Reductionism would be the negation of this claim. Views of testimonial knowledge are best seen as resultant belief theories. This is because, like resultant belief formulations of accounts of testimonial justification, they assess the epistemic status of the belief which results from testimony, not the transition from the belief that 'S said that p' to 'p'. One could maintain that an audience can justifiably make this transition without positive reasons to trust the speaker, and still maintain that testimonial knowledge requires positive reasons if one denied that audiences have a default right to trust their quasi-perceptions. Thus, the cases I give here will also count against many anti-reductionist views of testimonial knowledge.

out that a speaker said that *p* by being told by another speaker that they said that *p.*

We can divide up the resultant belief versions of reductionism and anti-reductionism in terms of the different claims which could be made about an agent's rights to trust in the credibility of one's informant, and the rights one has to trust in one's own understanding. Firstly, one could maintain that in order for an agent's testimonial belief to be justified it is neither necessary that the agent have reasons to trust in the credibility of the speaker, nor that they have reasons to trust in their own understanding. Such a view would endorse a default defeasible right to move from a quasi-perception of the speaker as having asserted that *p* to a belief that the speaker said that *p,* and a defeasible right to move from a belief that the speaker testified that *p* to a belief that *p.* Call this view 'strong anti-reductionism'. Strong anti-reductionism corresponds to transitional anti-reductionism with the additional claim that one has a defeasible default right to trust one's own understanding. It also corresponds closely to Tyler Burge's influential brand of anti-reductionsim[94]. Strong anti-reductionism is required if one wishes to endorse Burge's claim that it is possible to gain a priori testimonial knowledge.

Burge's (1993) argument for a priori testimonial knowledge runs roughly as follows: firstly he argues that perception plays a purely preservative role in understanding. That is, it introduces no empirical information into the audience's reasoning, merely triggering understanding of the content presented as true. He draws an analogy between testimony and memory: the function of memory in deductive reasoning is merely to recall propositions previously known, and present them to consciousness. No reasoning about memory itself is required. With testimony it is, according to Burge, as if we simply perceive what others have told us. Perception presents the message to consciousness without introducing any new content. Our perception of the speaker's utterance plays a merely causal role in our acquiring testimonial knowledge. We perceive and de-code the message without having to reason about the fact that the utterance was made, or the context in which it was made[95]. This part of Burge's argument corresponds closely to the claim that one has a default right to trust one's own understanding.

Secondly he argues that we have an a priori right to accept propositions presented to us as true. That is, if something is presented to us as true then we have a defeasible right to accept it. His argument for this latter claim is a little murkier, relating to the way a system capable of presenting propositions as true must be, in some way, systematically connected to the truth. This claim seems

---

[94] Burge's endorsement of the right to not only trust the speaker, but also to trust one's own understanding is discussed in Burge (1993, 1997), Christensen and Kornblith (1997), and Malmgren (2006).

[95] For objections to this part of Burge's argument see Bezuidenhout (1998), Christensen and Kornblith (1997), Longworth (2008), Malmgren (2006)

to correspond to the notion that we have a right to transition from being told that *p* to the belief that *p*. The two claims together seem to allow for a priori testimonial knowledge. When an a priori proposition is asserted the audience's perception will introduce no new subject matter, it will simply present the a priori proposition to the agent's consciousness. No a posteriori reasoning or reasons play a justificatory role here. Since the audience's entitlement to accept such presentations as true is also a priori, the resultant knowledge will be a priori.

Strong anti-reductionism will be my main target here. Thus, one upshot will be that we cannot achieve a priori testimonial knowledge. Not all brands of anti-reductionism will be rendered problematic. One could conceivably be a credulist regarding the transition from 'S said that p' to 'p', and still deny the default right hypothesis with respect to the output of one's own linguistic understanding. Call such a view 'weak anti-reductionism'. Transitional anti-reductionist views are neutral between strong and weak anti-reductionism. It might be initially unclear why one would endorse a default right hypothesis regarding trust in the speaker, but not trust in one's own understanding. After all, we regularly reason about the credibility of our informants, but rarely about our own understanding. However, it is not clear that the main motivations for anti-reductionism motivate strong anti-reductionism over weak anti-reductionism. For example, it is often claimed that reductionism is untenable since we frequently lack reasons to trust speakers even when we are intuitively warranted in doing so. However, it is less clear that we ever lack reasons to trust our own understanding. Moreover, arguments such as those presented by the assurance theorists concerning the way in which we see the speaker (as a source of evidence vs someone offering their word) only seem to support credulist views about the move from 'S said that p' to 'p'. They say nothing about our rights to trust our own understanding. Thus, for all I say here it is possible for one to maintain a weak anti-reductionist view[96].

Reductionist views can also be divided into strong and weak variants depending on the claims they make about the audience's right to trust their own understanding. A strong reductionist would claim that we have neither a default right to trust the speaker, nor such a right to trust our own understanding of the speaker. A weak reductionist would maintain that we have a default right to trust our understanding, but no default right to trust the speaker. It is not clear that weak reductionism is really reductionist at all, as it still involves a distinctive default right. That is, on such views one's right to accept testimony is not fully reduced to one's reasons to accept the

---

[96]   This is not to say that the arguments definitely don't carry over, simply that it is not clear whether or not they do. If the primary arguments for anti-reductionism support strong over weak anti-reductionism then the arguments of this chapter render anti-reductionism in general problematic.

testimony. Weak reductionist views will also be rendered problematic by the cases of cognitive penetrability.

Usually when views of testimonial justification are formulated it is done so in terms of the agent's rights with respect to *p* in cases where the speaker has said that *p*. This is true of both transitional and resultant belief formulations. That is, such views are normally formulated by reference to the 'good case': in cases in which the agent's understanding of the speaker is correct, does the agent need reasons to trust the speaker, and to trust their own understanding of the speaker, in order to be justified in believing that *p*? The views we have considered so far can be further divided on the basis of the claims they make about what we might call the 'bad case'. That is, they can be divided in terms of what they say about cases in which the audience misunderstands the speaker. In the good case the audience has a quasi-perception of the speaker as having said that *p*, and the speaker actually does say that *p*. In the bad case the audience has the same quasi-perception, but the speaker does not actually say that *p* - the quasi-perception is misleading. There are two stances a theorist could take to each type of case. Firstly one could maintain symmetry: agents have the same rights in the good case and the (phenomenologically indistinguishable) bad case. Secondly, one could endorse difference: despite the fact that the cases are phenomenologically indistinguishable agents have different rights in each case. My target here will be strong anti-reductionists and weak reductionists who endorse symmetry.

Strong reductionists and weak anti-reductionists are likely to endorse symmetry. After all, it is not clear why one would endorse a default right to trust one's understanding only when one has misunderstood the speaker. Moreover, internalists of all stripes are likely to endorse symmetry. This is because the good case and the bad case will be be phenomenologically indistinguishable to the agent, meaning that the epistemic rights they have in each case should (for the internalist) be the same. Internalist strong anti-reductionists and internalist weak reductionists will have problems with cases of cognitive penetrability.

It is open to externalists to deny symmetry. That is, it is open to externalists to claim that one has a default right to trust one's understanding only in cases in which one understands the speaker correctly. However, this not to say that it is easy for externalists to deny symmetry. Such a view is problematic for several reasons. Firstly, it is counter intuitive. Since the cases are phenomenologically indistinguishable to the agent it seems intuitive that the agent should have the same rights in each case. Of course, there are well known responses to this problem. The

externalist could simply maintain that the agent is rational or blameless for their belief in the bad case, and claim that this captures intuitions to the contrary[97]. More problematic is the fact that the externalist credulist is not exempt from providing an account of our right to trust our own understanding in the good case. In order to deny symmetry the externalist anti-reductionist must produce an account of our default right to trust our own understanding which does not predict that we have a default right to trust our understanding in the bad case. It is not clear quite how such a theory should go. For example, if one endorses the claim that we are entitled to trust our understanding in the good case because of the fact that our comprehension functions to produce true beliefs, then one should claim that we are also entitled to trust our understanding in the bad case, since we are still relying on a mechanism which functions to produce true beliefs[98]. Alternatively it might be claimed that our entitlement to rely on our understanding in the good case derives from the fact that such reliance on our own understanding is generally reliable (reliance on our understanding is a reliable way of gaining true beliefs about what is said). However, this also seems to carry over to the bad case. It is just as true in the good and bad cases that our linguistic understanding is generally reliable (which is not to say that, generally, our linguistic understanding is reliable in the good and bad case). Indeed, one could simply think of the general reliability claim as the hypothesis that there are significantly more good cases than bad cases. Thus, many externalist strong anti reductionists (and many externalist weak reductionists, if there are any such theorists) will also be forced to endorse symmetry.

My primary target will be strong anti-reductionists who endorse symmetry (call such theorists 'symmetrical strong anti-reductionists'). Strong anti-reductionism combined with symmetry entails that one has a defeasible default right to believe that *p* whenever one has a quasi-perception of a speaker as having said that *p*, regardless of whether or not the speaker actually did say that *p*. This will run into problems in some cases of miscommunication which arise from cognitively penetrated quasi-perceptions. In these cases a belief formed on the bases of the quasi-perception would be intuitively unjustified.

---

[97] In the cases discussed in the next section (especially the interpretation based cases) it appears that the agents involved may be blameworthy and irrational. So this strategy will not be completely straightforward. However, I believe that it is open to the externalist to claim that the agent is blameworthy and irrational in the cases presented here, yet blameless and rational in more standard cases of misunderstanding where the agent involved is not at fault in their misunderstanding.

[98] Burge (1993) can arguably be read this way, as can Graham (2010). For a similar view applied to perception see Schellenberg (2013)

3. Cognitively Penetrated Quasi-Perception.

In order to understand how quasi-perceptions can be cognitively penetrated, and to understand the relation between cognitive penetration of perceptual states and cognitive penetration of quasi-perceptual states, it is worth considering two distinct components of quasi-perception, both of which which might be influenced by background beliefs, attitudes, and associative mental states[99]. The contents of our quasi-perceptions are determined by our perceptions of certain sounds as words, and our assignment of meaning to those words (as well as other factors such as our grasp of underlying syntax). Let us consider speech perception first, as it is closer to Siegel's primary target and, as a form of perception, plausibly falls under the scope of her original arguments.

3.1. Speech Perception.

As discussed in the previous chapter, judgements about the beliefs, interests, and backgrounds of other interlocutors generate expectations concerning what they are likely to say, and thus about what words they are likely to use, and there is reason to suspect that these judgements and expectations influence speech perception. Firstly, if Siegel is correct about perception in general being cognitively penetrable then it follows that auditory perception should be cognitively penetrable[100] (although Siegel focuses on visual perception, she does take her arguments to extend to other types of perception). Moreover, as noted in chapter five, it is seemingly common to mistake one word for another on the basis of expectations about what a speaker is likely to say. Chapter five considered a case in which an audience member is in a noisy bar talking to what appears to be a business man, and mishears his utterance of 'I work on current opinions of markets' as 'I work on currency options markets'. This is an intuitively plausible case in which an agent's background beliefs and attitudes can have an impact on their speech perception, leading to a miscommunication. This case, although clearly epistemically problematic, doesn't seem to generate a problem for symmetrical strong anti-reductionists, as it seems plausible that the belief formed on the basis of this miscommunication is justified and rational. However,we can easily alter it so it falls in line with Siegel's examples:

---

[99]   I use the term 'components' loosely here. I do not mean to take a stand on the question of whether quasi-perceptions are composite states, or whether such a notion even makes sense.

[100] As in the case of visual perception it is ultimately an empirical question whether speech perception is cognitively penetrable.

INVESTMENT ADVICE: Sally has been tasked with headhunting an expert in currency swap options. However, she has been as yet unsuccessful, and is getting increasingly desperate. When she enters a bar and sees a man wearing jeans and a smart shirt she, purely through wishful thinking, forms the belief that he is an expert on currency options. As it happens he is actually a social scientist who works on public opinions of markets. Sally asks him what he works on and he says 'I work on current opinions of markets'. However, due to her expectations (based on her wishful thinking) she hears him as having said 'I work on currency options markets'. Sally takes this apparent testimony at face value, and takes it to confirm her prior belief that the man works on currency options.

This case seems no less plausible than Siegel's cases of cognitively penetrated visual perception. Sally has a quasi-perception of the man as having asserted that he works on currency options. Moreover, it does not look as if Sally possesses any defeaters concerning either her understanding, or the claim that the man works on currency options. The bar needn't be especially noisy, the assertion may well have sounded clear to Sally, and the man's appearance provides very little evidence concerning his occupation. The symmetrical strong anti-reductionist holds that if one has a quasi-perception as of the speaker saying *p*, and one has no reason to doubt the speaker, then one is justified in believing that *p*. Thus symmetrical strong anti-reductionism predicts that Sally's belief is justified. This is the wrong prediction. Sally's belief seems to have the same problematic features as Jill's belief that Jack is angry. It seems intuitively unjustified. It seems to involve an implicit move which, if explicit, would be judged to be straightforwardly irrational. And it seems circular, since her seemingly testimonial belief that the man is an expert on currency options depends on her prior belief that he is an expert on currency options. It seems that the symmetrical weak reductionist will make the wrong prediction here too. After all, Sally surely has most of the usual evidence we have of speaker reliability when talking to strangers. Thus, if she has a right to trust her understanding, and the case is parallel to other cases in which we have reasons to trust the testimony of a stranger, then she should, according to the symmetrical weak reductionist, be justified.

This case does seem problematic for the strong symmetrical anti-reductionist. However, it is not conclusive. Assuming that auditory perception is a-rational, it might be argued that once Sally has heard the speaker as uttering the words "I work on currency options markets" she has no rational choice but to accept that the speaker works on currency options. After all, she surely cannot be rationally criticised for hearing particular words. This would be like rationally criticising someone for having a hallucination. It might then be maintained that if Sally cannot be rationally criticised

then her belief is rational and/or justified. The viability of this response will depend on what one takes to be the relationship between rationality, justification, and blameworthiness. However, we needn't get into such issues here, as it will be argued in the following section that parallel arguments cannot be given in response to interpretational versions of cognitive penetrability cases. This is because, it will be argued, quasi-perceptions can, in some circumstances, be criticised as irrational. In cases of simple mishearing it would be inappropriate to criticise an agent for their defective quasi-perception. However, in cases of misinterpretation it will sometimes be appropriate to criticise agents for their quasi-perception.

We have seen that cases involving cognitively penetrated speech perception are inconclusive. Moreover, it is not clear that the type of cognitive penetration just discussed is interestingly different from the phenomena discussed by Siegel. After all, it just seems to involve cognitive penetration of our perception of the audible sounds produced by a speaker. That this is a possibility follows straightforwardly from Siegel's general claim that perception is cognitively penetrable. This result will be completely unsurprising to those who already accept Siegel's arguments, and it will be unpersuasive for those who already reject them. Thus it is worth considering a second way in which quasi-perception can be cognitively penetrated.

3.2. Interpretation.

It has been a theme of the previous chapters that much of language is context sensitive, and that even semantically context insensitive assertions are often interpreted in a way that is sensitive to the needs of the context. For example, consider the term 'vulnerable'. This is not standardly seen as a context sensitive term. However, as we saw in chapter five, it can be used and understood in subtly different ways in different contexts. Consider two speakers, one is a member of a charity which provides blankets, food, and heating for old people living in poverty, the other is a member of a criminal gang who brazenly steals from old people who are unable to put up a fight. An utterance of 'Alice is vulnerable' will mean subtly (but importantly) different things in the mouths of each speaker (at least, in contexts where they are planning their respective activities). For one 'vulnerable' will mean 'at risk from hunger and the cold' (this reading will be denoted as 'vulnerable$_1$'), for the other it will mean 'an easy/helpless target' (this will be denoted as 'vulnerable$_2$'[101]). Thus, the way we understand utterances containing the word 'vulnerable' will

---

[101] It is important to note that these two readings are genuinely distinct. There are people who may fall in the extension

depend on the way we modulate the term (and whether or not we do modulate it) in response to the needs of the context. Since we rely on our representation of the context in assigning meaning, it should be clear that misrepresentations of the context can lead to misinterpretation. For example, if one falsely believed that the speaker of 'Alice is vulnerable' were a member of the criminal gang planning a robbery, then one would likely misinterpret the speaker as saying that Alice is vulnerable$_2$.

Our quasi-perceptions of utterances are perception like experiences of utterances as having a particular meaning. The phenomenology of context sensitive or modulated speech perception is no different to that of context insensitive speech perception (if there is such a thing). Thus, the background representations which determine our modulations or assignment of meanings to context sensitive terms affect the contents of our quasi-perceptions. This would be the case even if our actual perception of the words spoken (and perception more generally) were cognitively impenetrable. The influence of background representations of the speaker on our assignments of meaning can generate cases in which justification and rationality is undermined. The case of Tom Robinson discussed in chapter five provides a plausible example of this. Indeed, we can conceive a case based on that of Tom Robinson which is directly analogous to Siegel's examples:

> RACIST JACK: Tom, a black man, has been falsely accused of stealing from Alice, an old woman living in poverty. Jack, a racist, is on the jury. As it happens Tom was simply in the wrong place at the wrong time. He was at Alice's home intending to offer her help. However, a robbery was in progress as he entered the premises. At the start of the trial Jack forms the belief that Tom is guilty for no other reason than the colour of his skin. He congers up a narrative in which Tom had been scheming to steal from poor defenceless Alice. So when Tom is asked why he was at Alice's residence, and he replies 'because she was vulnerable' Jack naturally interprets Tom as having revealed that he saw Alice as an easy target. He takes this to confirm his prior belief that Tom saw Alice as an easy victim, and that he was in fact the robber.

In this case Jack quasi-perceives Tom as saying that he saw Alice as vulnerable$_2$. We can imagine that he possessed no defeaters concerning his own understanding or Tom's actual attitude[102]. Thus, he is, according to symmetrical strong anti-reductionism, justified in believing on the basis of his

---

of vulnerable$_1$ but not vulnerable$_2$. For example, a poor old lady trained in krav maga.

[102] We can imagine that this took place at the start of the trial when Jack had very little information concerning Tom's innocence.

quasi-perception that Tom saw Alice as vulnerable$_2$. His belief here seems clearly unjustified. And it also seems viciously circular, as it is reliant on his previous belief that Tom saw Alice as vulnerable$_2$, a belief he formed purely on the basis of Tom's race. Thus, symmetrical strong anti-reductionism is once again rendered untenable. Moreover, this conclusion holds even if Siegel is wrong about the cognitive penetrability of perceptual states. Thus the case against symmetrical strong anti-reductionism seems stronger than the case against perceptual dogmatism.

In the previous section it was noted that the symmetrical strong anti-reductionist might bite the bullet and maintain that Sally was justified. The reasoning was that Sally cannot be blamed or judged irrational for mishearing the speaker. A similar worry may arise here. That is, one might still worry that, since Jack possesses no defeaters, he cannot be rationally criticised for trusting his understanding, and is therefore not rationally criticisable for his belief. This would be a mistake. Unlike the case of perceptual seemings (including penetration of speech perception) which are arguably a-rational, we can be rationally criticised for our quasi-perceptions. For example, people who's quasi-perceptions fail to track contextual evidence seem to be rationally criticisable. Suppose you point very clearly at an apple and say 'that is my apple'. If I were to see your gesture and yet still misunderstand you and take you to be talking about a different apple, despite your very clear gesture, you would be within your rights to criticise me. This is not the case for faulty perceptual seemings. For example, if I hallucinated, or merely had some strange preoccupation with a particular apple which caused me to actually perceive you as pointing to that apple then you would not be able to hold me accountable for any epistemic wrongdoing. But if I clearly saw you point at the intended apple you would be able to hold me responsible. Indeed, I would surely be irrational if I were to see your clear gesture and yet still take you to be talking about a different, contextually non-salient apple. The upshot of this is that even if Jack can't be criticised for the move from understanding to belief he can be criticised for his understanding. Thus his over all belief is rendered criticisably irrational. Indeed, the notion that Sally might be justified, whilst Jack is unjustified, seems to pattern well with intuitions about the cases. Intuitively Jack's belief is epistemically worse than Sally's, and this fact cries out for explanation[103]. An obvious explanation for this is that Jack's belief, unlike Sally's is unjustified.

The notion that quasi-perceptions can be judged rational or irrational also serves to illustrate why an

---

[103] If it is not initially clear that Jack's mistake is worse than Sally's consider the fact that Sally at least appeared to hear the words "I work on currency options markets". It is practically unheard of to criticise someone for mishearing an utterance. It is, however, fairly common to criticise audiences for misinterpreting speakers. We will return to this point shortly.

important response to Siegel's arguments against perceptual dogmatism does not carry over to the cases presented here. Siegel (2013) discusses several features which separate perceptual seemings from beliefs, and argues that these features cannot be the features which allow beliefs to transmit their ill-foundedness, meaning that whatever features do allow beliefs to transmit ill-foundedness are shared with perceptual seemings. Richard Fumerton (2013) responds, arguing that the rational assessability of beliefs accounts for their ability to transmit irrationality. Perceptual seemings may have rationally assessable etiologies, but they are not themselves rationally assessable. Thus, perceptual seemings do not share with beliefs the feature which allows beliefs to transmit ill-foundedness. This objection does not carry over to the testimonial case since quasi-perceptions, as we have just seen, do seem to be rationally assessable. Thus, if Fumerton is correct, quasi-perceptions share with beliefs the features which allow beliefs to transmit ill-foundedness.

The claim that quasi-perceptions can, in some cases, be rationally criticisable will be buttressed in the next section where, in response to Jack Lyons's response to Siegel's argument, it is argued that quasi-perceptions can be based.

4. Response: Circularity and Basing.

Perceptual seemings are importantly different from beliefs, and the differences between the two have been a source of several objections to Siegel's arguments. One such difference, discussed in the previous section, is the ability of beliefs, and inability of perceptual seemings to be irrational. A second important (and related) difference is the ability of beliefs to be based. It is argued in this section that quasi-perceptions, are not only rationally criticisable, but are also capable of being based.

Jack Lyons (2011) argues that the problem with cognitive penetrability cannot be that it allows for viciously circular beliefs to be justified[104]. The reason for this is that circularity is a matter of improper basing. That is, epistemic circularity occurs when an agent's only evidence for a belief B_1 is either a prior belief in B_1, or is epistemically dependent on the agent's prior belief in B_1. However, as Lyons points out, perceptual seemings to not not appear to be the sort of things which

---

[104] Lyons does not claim that the agents in Siegel's cases are justified, rather he claims that they are unjustified because their beliefs are unreliably formed. He takes Siegel's cases to motivate a reliablist view of justification. I do not wish to commit to a reliablist view of justification. Thus I wish to avoid accounting for the lack of justification in these cases in terms of reliability. Moreover, as noted in section two, I think it is plausible that many reliablist theories of our right to trust our understanding will also make the wrong predictions about these cases.

are 'based'. You cannot receive evidence for or against a perceptual seeming, they are not responsive to reasons. It does not make sense to ask someone why they see a banana as yellow for example. The answer will simply be 'because it is yellow', or 'it just looks yellow'. The existence of cognitive penetration appears to show that perceptual seemings can be causally based on prior beliefs. But this is not sufficient for circularity, circularity requires epistemic evidence responsiveness.

Indeed, one of the guiding thoughts in Siegel's (2013) development of her argument is that beliefs based on checkered experiences may not be based on the checkering process. She points out that if a belief B_1 is based on another belief B_2 then a change in B_2 will usually trigger a change in B_1. However, we rarely adjust beliefs based on a checkered experiences in response to changes in the checkering state. For example, if one were petrified of guns, and this fear caused one to see something in one's fridge as a gun, then one would not lose one's belief that there was a gun in the fridge upon losing one's fear of guns. You would still think that you saw a gun. Siegel's (2013) formulation of the problem moves away from the circularity claim and avoids any suggestion that the problematiic beliefs are based on the checkering process. Instead she considers ways in which checkered experiences differ from beliefs, and argues that none of these differences explain why beliefs, but not perceptual seemings, are able to transmit ill-foundedness.

It might be thought that these worries extend to the testimonial version of the problem. That is, it might appear that quasi-perceptions are not the kind of things which can be based either. After all, the phenomenology of quasi-perception is similar to that of normal perception. When we explicitly reason regarding the meaning of someone's utterance it might be thought that the resultant state of understanding is not really a quasi-perception (after all, one of the distinguishing features of quasi-perceptions was their phenomenology, and the resultant state here may have a markedly different phenomenology).

Before responding to this worry it is worth putting to one side a bad reason to think that quasi perceptions can be based: Some theorists maintain that states of linguistic understanding are simply states of belief (or knowledge). For example, some cognitivist accounts of linguistic understanding hold that such understanding consists in, or requires, belief or knowledge of a compositional semantic and syntactic theory[105]. Thus, it might be thought that states of linguistic understanding, being beliefs, can obviously be based. This line of argument is unconvincing. Even if the cognitivist

---

[105] Pettit (2002) argues, on epistemic grounds, that linguistic understanding does not require knowledge or belief concerning meaning. Gross (2005) defends the belief hypothesis against Pettit's attack.

is correct that states of understanding are states of belief, such beliefs are purely tacit. Thus, it is an open question whether they are genuinely evidence responsive like normal beliefs. Moreover, the perception-like phenomenology of quasi-perceptions gives us good reason to distinguish them from beliefs. However, beliefs are not the only things which are susceptible to basing. For example, fears or suspicions can be justified or unjustified, and based on prior beliefs and attitudes. So the mere fact that quasi-perceptions are not beliefs does not give us reason to conclude that they cannot be based. Moreover, there are important differences between perceptual seemings and quasi-perceptions which lend credence to the hypothesis that quasi-perceptions can be based.

Siegel's main reason for avoiding the claim that checkered beliefs are based on the checkering process is that a change in the penetrating state will not usually trigger a change in the checkered belief. This claim has less plausibility in the case of checkered testimonial belief. We experience miscommunication often, and we are accustomed to adjusting our beliefs about what we have been told when we adjust our conception of the speaker (and their character, goals, and knowledge etc.). Sometimes these reassessments are explicit, other times they merely affect the way that we recall a situation. For example, suppose that racist Jack has an epiphany and realises that his racist views were wrong all along. This is likely to cause him to see many of his past interactions with black people in a new light. His quasi-perception of Tom as saying that Alice was vulnerable$_2$ was penetrated by an picture he had built up of Tom as a predatory scheming criminal. Once this conception is reversed Jack is likely to see Tom's utterance in a new light, and revise his belief. This revisability in response to altered conceptions of the social situation or speaker is a feature which testimonial beliefs do not seem to share with perceptual beliefs. These considerations seem to suggest that in the case of testimony the resultant beliefs may be partly based on the checkering process.

Additionally, it seems that quasi-perceptions, unlike perceptual seemings, can sometimes be retrospectively rationalised and justified. For example, if someone points to an apple and says 'that is red' you will have a quasi-perception of their utterance as meaning that the particular apple they are pointing at is red. If asked why you understood the speaker as saying that the apple was red you would be able to quickly respond 'because I saw her pointing at the apple'. This suggests that your quasi-perception was evidence responsive. That is, it was partially based on your perceptual experience of the speaker pointing to a particular apple. This apparent basing seems to be

accessible to consciousness, even though the phenomenology is non-inferential[106]. Indeed, this ability of quasi-perceptions to be retrospectively rationalised not only supports the hypothesis that they can be based, but also that they are capable of being rational or irrational. Moreover, the rational flaw in the cases of seemingly irrational quasi-perception (where the speaker clearly points to a particular object, yet a different referent is assigned) seems to be that the quasi-perception is not properly evidence responsive (and thus badly based). Thus, the cases discussed at the end of the previous section also seem to support the hypothesis that quasi-perceptions can be based. As a result, Lyons's argument against the circularity formulation of Siegel's argument does not carry over to the testimonial variant of the argument.

5. *Response: Defeaters and Monitoring.*

I have argued that symmetrical strong anti-reductionism makes the wrong predictions about cases in which the audience inexcusably misinterprets the speaker and forms an unjustified belief. The symmetrical strong anti-reductionist might claim that they do not actually make the problematic predictions, since most forms of anti-reductionism contain a no defeater clause, and there are defeaters present in the cases presented above. This thought could be spelled out different ways. It could be spelled out in terms of subpersonal monitoring for signs of unreliability, or it could be spelled out in terms of the agents in the cases actually possessing defeaters to the claim that their understanding is reliable.

The monitoring theorist will maintain that although audiences need not possess positive reasons to trust the speaker or their own understanding, they must nonetheless be sensitive to signs that their understanding (or the speaker) is unreliable. It is not clear that such a view is really a form of anti-reductionism. Indeed, it seems to be a version of Elizabeth Fricker's local reductionism. However, it might be thought that if the monitoring is subpersonal then it does not constitute part of the audience's agent level reasoning. Thus it might be thought that such subpersonal monitoring is compatible with anti-reductionism (see Goldberg and Henderson (2006) for a view along these lines, see Fricker (2006b) for a response). If subpersonal monitoring is compatible with anti-

---

[106] Indeed, considerations along these lines have been taken to show that linguistic understanding consists in inferential propositional knowledge (for example, Stanley (2005)). Stanley argues that linguistic understanding is always based on contextual knowledge because such knowledge is required in order for us to tell whether or not an assertion violates the gricean maxim of manner, and thus whether or not it should be interpreted literally. However, even if we do not accept such an extreme conclusion, these considerations still seem strongly suggestive of the hypothesis that states of quasi-perception, despite their phenomenology, can be based.

reductionism then it might be claimed that we monitor for signs that we have misinterpreted the speaker. We might, for example, be sensitive to indicators such as hearing exactly what we wanted to hear, which indicate that wishful thinking has played a part in interpretation.

The monitoring response is problematic. It might provide a plausible account of the Sally case, however not all cases of problematically penetrated quasi-perception involve deviant causes such as wishful thinking. Some merely involve background mental states (such as attitudes and beliefs) influencing interpretation the same way they do in the good case. For example, in cases of semantic context sensitivity or modulation we rely on our representation of the speaker in determining the correct interpretation. Many of the problem cases arise when these background representations of the context are themselves irrationally formed. The actual process by which these representations are drawn upon in order to determine an interpretation are the same across both the good and bad cases (this seems to be the case with Racist Jack). Thus in many cases there is nothing unusual about the quasi-perception for the sub-personal monitoring mechanism to pick up on (at least, nothing that it would not pick up on in the good case as well).

A second version of the defeater response holds that the agents in the cases I have presented possess positive defeaters for their beliefs. That is, it might be thought that the majority of adult speakers have enough experience to know the sorts of situations in which they are likely to misinterpret people. The plausibility of this response will depend on the defeaters which are postulated. Once again the response seems to work better for Sally than for Jack. Perhaps any agent in Sally's situation could be expected to be weary of their interpretation. After all, we all know that we sometimes misinterpret people in high stakes situations, especially when we really want to hear a particular thing. It is harder to see what defeater Jack might possess. For example, it is unlikely that he possesses any defeaters concerning his ability to interpret black people in an unbiased way. I think that the most promising way to develop this objection is to hold that there are certain very general defeaters possessed by all adult speakers which serve to block default trust in one's quasi-perceptions in almost all cases. For example, every adult speaker is surely aware of the fact that we sometimes misinterpret one-another. Perhaps this serves as a defeater to Jack's reliance on his understanding. Of course this defeater is also present in the good cases, so we would have to maintain that the default right of adult speakers to rely on their understanding is always defeated. However, this might not be seen as a problem. After all, it might be thought that adults possess enough positive and negative information that they do not need to rely on default rights to achieve testimonial knowledge. On such a view anti-reductionism and reductionism will only differ in their

predictions about the testimonial beliefs of small children.

Unfortunately, even the general defeater approach will not vindicate symmetrical strong anti-reductionism. This is because we can produce cases of problematic cognitive penetration involving small children who are not yet experienced enough to possess general beliefs about miscommunication. Consider the following case:

> MARBLES: Tommy, a small child, has a collection of marbles. He is very proud of his marbles, and he thinks of them constantly. As a result, whenever anyone talks about marbles he takes them to be talking about his marbles (even when he has no reason to think they would know about his marbles). One day at playgroup the playgroup leader says 'the marbles are very special', intending to refer to the marbles available for the children at the play group. However, Tommy misinterprets the playgroup leader and takes her to be talking about his marbles (despite the fact that he has no reason to think that he knows about his marbles). Indeed, he does so partly on the basis of his prior belief that his marbles are very special. Tommy thus reaffirms his belief that his marbles are special.

In this case Tommy has a quasi-perception of the playgroup leader as saying that his marbles are very special. Symmetrical strong anti-reductionism holds that if one has a quasi-perception of a speaker as saying *p* and one has no reason to either distrust the speaker or one's own understanding, then one is justified in believing that *p*. Tommy does not possess any defeaters. He has no positive reason to distrust his play group instructor (indeed, we might maintain that she has always been a highly reliable informant, as people often are with children). Moreover, Tommy is too young to possess any defeaters concerning the general reliability of his ability to interpret others. He might not even be aware of the phenomenon of misinterpretation. Thus, symmetrical strong anti-reductionism predicts that Tommy's belief is justified. However, his belief is surely not justified. Indeed, it seems problematic in much the same way as racist Jack's belief (that is, it seems *epistemically* problematic in the same way. Jack's belief was also morally problematic). Thus, symmetrical strong anti-reductionism still makes the wrong prediction.

6. Conclusion.

I have outlined a series of cases in which symmetrical strong anti-reductionism appears to make the

wrong predictions about whether or not an agent's belief is justified. These cases all involve cognitively penetrated quasi-perceptions, and are analogous to the cases of cognitively penetrated perception presented by Susanna Siegel in response to perceptual dogmatism. Whilst discussing these cases I illustrated several important similarities and dissimilarities between quasi-perception, beliefs, and perceptual seemings, and used these analogies to argue that the case against symmetrical strong anti-reductionism is stronger than the case against perceptual dogmatism. Strong anti-reductionism is required to support the claim that there is a priori testimonial knowledge. Thus, if one wishes to endorse a priori testimonial knowledge one must endorse non-symmetrical strong anti-reductionism. In order to do so it is necessary to produce an account of our default right to trust in our understanding in the good case which does not carry over to the bad case. I have not argued that such an account cannot be given. However, it is unclear how it would go, and the onus is on the proponent of a priori testimonial knowledge to give such an account.

Bibliography.

Åkerman, J. 2009. A Plea For Pragmatics. *Synthese*. 170. 155-167.

Åkerman, J. 2010. Communication and Indexical Reference. *Philosophical Studies*. 149. 355-366.

Alicke, D. 1993. Egocentric standards of conduct evaluation. *Basic and Applied Social Psychology*. 14. 171–192.

Alston, W. 2005. Religious Language. From Wainwright (Ed), *The Oxford Handbook of Religious Language* (pp 220-242). Oxford university Press.

Baumann, P. Forthcoming . No Luck With Knowledge? On a Dogma of Epistemology. *Philosophy and Phenomenological Research*. 1-29.

Bezuidenhout, A. 1997. The Communication of De Re Thoughts. *Nous* 31 (2): 197-225.

Bezuidenhout, A. 1997. Pragmatics, Semantic Underdetermination and the Referential/Attributive Distinction. *Mind*. 106. 375-409.

Bezuidenhout, A. 1998. Is Verbal Communication a Purely Preservative Process. *Philosophical Review.* 107 (2). 261-288.

Bezuidenhout, A. 2001. Metaphor and What is Said: A Defence of the Direct Expression View of Metaphor. *Midwest Studies in Philosophy.* 25. 156-186.

Bezuidenhout, A. 2002. Truth Conditional Pragmatics. *Philosophical Perspectives* 16: 105-134.

Bloom, P. (2000). How Children Learn the Meaning of Words. Cambridge Mass: MIT Press/Bradford Books.

Bogardus, T. 2014. Knowledge Under Threat. *Philosophy and Phenomenological Research.* 88. (2): 289-313.

Boland, J., Clark, S. Manuscript. Homophone Disambiguation and Vocal Stereotypes.

Bolinger, D. 1968. Aspects of Language. New York: Harcourt, Brace and World.

BonJour, L. 1980. Externalist Theories of Empirical Knowledge. *Midwest Studies in Philosophy.* 5. 53–73.

Borg, E. 2004. Minimal Semantics. Oxford University Press.

Borg, E. 2012. Pursuing Meaning. Oxford University Press.

Bratman, M. 1992. Practical Reasoning and Acceptance in Context. *Mind.* 101. (401). 1-15.

Brooks, A., Huang, L., Kearney, A., Murray, F. Forthcoming. Investors Prefer Entrepreneurial Ventures Pitched by Attractive Men. *Proceedings of the National Academy of Sciences of the United States of America.* p 1-5.

Buchanan, R. 2010. A Puzzle about Meaning and Communication. *Nous.* 44 (2). 340-371.

Buchanan, R. 2012. Is Belief a Propositional Attitude. *Philosophers Imprint* 12 (1). 1 - 20.

Buckwalter, W., Rose, D., Turri, J. Forthcoming. Belief through Thick and Thin. *Nous.* 1-28.

Burge, T. 1993. Content Preservation. *Philosophical Review.* 102 (4). 457-488.

Burge, T. 1999. Comprehension and Interpretation. From L.Hahn (Ed), *The Philosophy of Donald Davidson,* 229-254. Chicago: Open Court Publishers.

Burge, T. 2010. The Origins of Objectivity. Oxford University Press.

Burge, T. 2013. Postscript: Content Preservation. From Burge. T, *Cognition Through Understanding*, 254-285. Oxford University Press.

Campbell-Kibler, C. 2010. Sociolinguistics and Perception. *Language and Linguistics Compass.* 4 (6). 377-389

Cappelen, H, Lepore E. Indexicality, Binding and A Priori Truth. *Analysis*. 62. 272-281.

Cappelen, H. and Lepore, E. 2004. Insensitive Semantics: A Defence of Semantic Minimalism and Speech Act Pluralism. Blackwell Publishers.

Cappelen, H. and Lepore, E. 2006. Shared Content. from Lepore & Smith (eds) *The Oxford Handbook of Philosophy of Language,* 1056- 1069. Oxford University Press.

Casscells, W., Schoenberger, A., Grayboys, T. 1978. Interpretation by Physicians of Clinical Laboratory Results. *New England Journal of Medicine*. 299. 999-1000

Carston, R. 2002a. Thoughts and Utterances: The Pragmatics of Explicit Communication. Wiley Blackwell.

Carston, R. 2002b. Metaphor, Ad Hoc Concepts, and Word Meaning: More Questions than Answers. *UCL Working Papers in Linguistics*. 14. 83-105.

Carston, R., Wilson, D. 2007. A Unitary Approach to Lexical Pragmatics: Relevance, Inference, and Ad Hoc Concepts. In Burton-Roberts (Ed), *Pragmatics,* 230-259. Basingstoke and New York Palgrave Macmillan.

Casasanto, L. 2008. Does Social Information Influence Sentence Processing? *30th Annual meeting of the Cognitive Science Society.* 1-6

Christensen, D., Kornblith, H. 1997. Testimony, Memory, and the Limits of the A Priori. *Philosophical Studies*. 86. 1-20.

Clapp, L. (2012). Three Challenges for Indexicalism. *Mind and Language*. 27 (4). 235-465.

Clark, H., Krych, A. (2001). Speaking While Monitoring Addressees for Understanding. *Journal of Memory and Language*. 50. 62-81.

Coady, T, 1992. Testimony: A Philosophical Study. Oxford University Press.

Cohen, J. 2013. Indexicality and the the puzzle of the answering machine. *The Journal of Philosophy*. 110. (1). 5-32.

Comesana, J. 2005. Unsafe Knowledge. *Synthese*. 146: 395-404.

Converse, A., Lin, S., Keysar, B., and Epley, N. 2008. In the Mood to Get Over Yourself: Mood Affects Theory of Mind Use. *Emotion*. 8 (5). 725-730.

Corazza, E., Fish, W., Gorvett, J. 2002. Who Is I?. *Philosophical Studies*. 107. 1-21.

Cosmides, L., Tooby, J. 1996. Are Humans Good Statisticians After All? Rethinking Some Conclusions from the Literature on Judgement Under Uncertainty. *Cognition*. 58 (1). 1-73.

Creel, S., Bergman, M. 2011. How Talker Identity Relates to Language Processing. *Language and Linguistics Compass*. 5 (5). 190 - 204.

Davidson, D. 1973. Radical Interpretation. *Dialectica*. 27. 314-328.

Donnellan, K. 1966. Reference and Definite Descriptions. *Philosophical Review* 75 (3): 281-304.

Donnellan, K. 1968. Putting Humpty Dumpty Together Again. *Philosophical Review*. 77 (2). 203-215.

Dorr, C., Hawthorne, J. 2014. Semantic Plasticity and Speech Reports. *Philosophical Review*. 123 (3): 281-338.

Dunning, D., Meyerowitz, J. A., Holzberg, A. 1989. Ambiguity and self-evaluation: The role of idiosyncratic trait definitions in self-serving assessments of ability. *Journal of Personality and Social Psychology.* 57. 1082–1090.

Epley, N., Savitsky, K., & Gilovich, T. 2001. Empathy neglect: Reconciling the spotlight effect and correspondence bias. *Journal of Personality and Social Psychology.* 83. 300–312.

Epley, N., Morewedge, C., Keysar, B. 2004. Perspective Taking in Children: Equivelant Egocentrism but Differential Correction. *Journal of Experimental Social Psychology.* 40. 760-768.

Faulkner, Paul. 1998. David Hume's Reductionist Epistemology of Testimony. *Pacific Philosophical Quarterly.* 79. 302-313.

Faulkner, Paul. 2011. Knowledge on Trust. Oxford University Press.

Fricker, E. 1994. Against Gullibility. In B. K. Matilal and A. Chakrabarti (Ed.), *Knowing from Words,* 125-161. Amsterdam: Kluwer Academic Publishers.

Fricker, E. 1995. Telling and Trusting: Reductionism and Anti-Reductionism in the Epistemology of Testimony. *Mind* 104: 393-411.

Fricker, E. 2003. Understanding and Knowledge of What is Said. From A Barber (Ed) *Epistemology of Language,* 325-367. Oxford University Press.

Fricker, E. 2006a. Second Hand Knowledge. *Philosophy and Phenomenological Research.* 73. (3). 592-618.

Fricker, E. 2006b. Varieties of Anti-Reductionism About Testimony: A Response to Goldberg and Henderson. *Philosophy and Phenomenological Research.* 72. (3). 618-628.

Fricker, E. 2012. Stating and Insinuating. *Proceedings of the Aristotelian Society - Supplementary Volume.* 86. 61-95.

Fricker, M. 2007. Epistemic Injustice: Power and Ethics of Knowing. Oxford University Press.

Fumerton, R. 2013. Siegel on the epistemic impact of 'checkered' experience. *Philosophical Studies.* 126. 733-739.

Geiselman, R., Bellezza, F. 1977. Incidental Retention of Speaker's Voice. *Memory and Cognition.* 5 (6). 658-665.

Gendler, T, S. 2011. On the Epistemic Cost of Implicit Bias. *Philosophical Studies.* 156 (1). 33-63.

Gilbert, D. 1991. How Mental Systems Believe. *American Psychologist.* 46. (2). 107-119.

Gilovich, T., Medvec, V. H., & Savitsky, K. 2000. The spotlight effect in social judgment: An egocentric bias in estimates of the salience of one's own actions and appearance. *Journal of Personality and Social Psychology.* 78. p 211–222.

Goldberg, S. 2001. Testimonially based Knowledge from False Testimony. *Philosophical Quarterly* 51:205. 512-26

Goldberg, S. 2005. Testimonial Knowledge from Unsafe Testimony. *Analysis.* 65 (4). 302-311.

Goldberg, S. 2006. Reductionism and the Distinctiveness of Testimonial Knowledge. From Lackey and Sosa (Eds), *The Epistemology of Testimony,* 127-144. Oxford University Press.

Goldberg, S. 2007. Anti-Individualism - Mind and Language, Knowledge and Justification. Cambridge University Press.

Goldberg, S. 2011. Putting the Norm of Assertion to Work: The case of Testimony. From Cappelen and Brown (Eds), *Assertion: New Philosophical Essays,* 175-197. Oxford University Press.

Goldberg, S. 2013. Anonymous Assertions. *Episteme.* 10. (2). 135-151.

Goldberg, S., Henderson D. 2006. Monitoring and Anti-Reductionism in the Epistemology of Testimony. *Philosophy and Phenomenological Research.* 72 (3). 576-593.

Goldman, A. 1976. Discrimination and Perceptual Knowledge. *Journal of Philosophy.* 73. (20). 771-791.

Graham, P. 2010. Testimonial Entitlement and the Function of Comprehension. From Pritchard, Duncan., Miller Alan., and Haddock, Adrian. (Eds), *Social Epistemology,* 148-174. Oxford University Press.

Graham, P. 2012. Epistemic Entitlement. *Nous*. 46 (3). 449-482.

Grice, H.P. 1989. Studies in the Way of Words. Cambridge MA: Harvard University Press.

Gross, S. 2005. Linguistic Understanding and Belief. *Mind.* 114. 453. 61-66.

Gorvett, J. 2005. Back Through the Looking Glass: On the Relationship Between Intentions and Indexicals. *Philosophical Studies*. 124. 295-312.

Hall, A. 2008. Free Enrichment or Hidden Indexicals. *Mind and Language*. 23. (4). 426-456.

Hanna, J., Tanenhaus, M. & Trueswell, J. 2003. The effects of common ground and perspective on domains of referential interpretation. *Journal of Memory and Language*. 49. 43–61.

Hawley, K. 2014. Trust, Distrust, and Commitment. *Nous*. 48 (1). 1-20.

He, Z., Bolz, M., Baillargeon, R. 2012. False-belief understanding in 2.5-year-olds: evidence from violation-of-expectation change-of-location and unexpected-contents tasks. *Developmental* Science 14 (2). 292-305

Heck, R. 2002. Do Demonstratives Have Senses?. *Philosopher's Imprint* 2 (2). 1-33.

Hinchman, E. 2005. Telling as Inviting to Trust. *Philosophy and Phenomenological Research*. 70.

(3). 562-587.

Hornsby, J. Langton, R. 1998. Free Speech and Illocution. *Legal Theory*. 4. p 21-37.

Horton, W., Keysar, B. 1996. When do Speakers Take Into Account Common Ground? *Cognition*. 59. 91-117.

Jenkins, C. 2007. Entitlement and Rationality. *Synthese*. 157. 25-45.

Jones, K. 2002. The Politics of Credibility. From Antony and Charlotte (Eds), *A Mind of One's Own: Feminist Essays on Reason and Objectivity, 2nd edition,* 154-176. Boulder: Westview.

Kaplan, David. 1989. Afterthoughts. From Perry and Wettstein (Eds), *Themes from Kaplan*. Oxford University Press.

Kaplan, M. 1996. Decision Theory as Philosophy. Cambridge University Press.

Kelp, C. 2009. Knowledge and Safety. *Journal of Philosophical Research.* 34: 21-31.

Keren, A. 2012. On the Alleged Perversity of the Evidential View of Testimony. *Analysis*. 72. (4). 700-707.

Keysar, B. 1994. The Illusory Transparency of Intention: Linguistic Perspective Taking in Text. *Cognitive Psychology*. 26. 165-208.

Keysar, B., Barr, D., Balin, J., Paek, T. 1996. Definite Reference and Mutual Knowledge: Processing Models of Common Ground in Comprehension. *Journal of Memory and Language*. 39. 1-20.

Keysar, B., Barr, D., Balin, J., and Brauner, J. (1996). Common Ground: An Error Correction Mechanism in Comprehension. *Annual Meeting of the Psychonomic Society, Chicago* (reported in Keysar, Barr, and Horton 1998).

Keysar, B., Barr, D., Horton, W. 1998. The Egocentric Basis of Language Use: Insights from a Processing Approach. *Current Directions in Psychological Science*. 47-50.

Keysar, B. 2007. Communication and miscommunication: the role of egocentric processes. *Intercultural Pragmatics*. 4. 71-84.

Keysar, B., Henly, A. 2002. Speakers' Overestimation of their Effectiveness. *Psychological Science*. 13 (3). 207-212.

King, J. 2013. Supplementives, the Coordination Account, And Conflicting Intentions. *Philosophical Perspectives*. 27. p 288-311.

King, J. 2014. Speaker Intentions in Context. *Noûs* 48 (2). 219-237.

King, J. Forthcoming. The Metasemantics of Contextual Sensitivity. From A. Burgress and B. Sherman (Eds) *New Essays in Metasemantics*. Oxford University Press.

Kukla, R. Forthcoming. Performative Force, Convention, and Discursive Injustice. *Hypatia*. 1-18.

Lackey, J. 1999. Testimonial Knowledge and Transmission. *The Philosophical Quarterly*. 49. 471-490.

Lackey, J. 2003. A Minimal Expression of Non-Reductionism in the Epistemology of Testimony. *Nous*. 37. 706-723.

Lackey, J. 2006. It Takes Two to Tango: Beyond Reductionism and Anti-Reductionism in the Epistemology of Testimony. From Lackey and Sosa (eds.),*The Epistemology of Testimony,* 160-189. Oxford University Press.

Lackey, J. 2008. Learning from Words: Testimony as a Source of Knowledge. Oxford University Press.

Lee, H. 1960. To Kill a Mocking Bird. J.B. Lippincott & Co.

Lin, S., Keysar, B., and Epley, N. (2010). Reflexively Mindblind: Using Theory of Mind to Interpret Behaviour Requires Effortful Attention. *Journal of Experimental Social Psychology*. 46. p 551-556.

Longworth, G. 2008. Comprehending Speech. *Philosophical Perspectives*. 22. 339-373.

Ludlow, P. 2014. Living Words: Meaning Underdetermination and the Dynamic Lexicon. Oxford University Press.

Luzzi, F. 2010. Counter-Closure. *Australasian Journal of Philosophy* 88 (4). 673-683

Lyons, J. 2011. Circularity, Reliability, and the Cognitive Penetrability of Perception. *Philosophical Issues*. 21. 289-311.

MacFarlane, J. 2005. Making Sense of Relative Truth. *Proceedings of the Aristotelian Society*. 105. 321-339.

MacFarlane. 2011. What is Assertion?. From Brown and Cappelen (Eds), *Assertion,* 79-96. Oxford University Press.

Maitra, I. 2007. How and Why to be a Moderate Contextualist. From Preyer and Peter (Eds), *Context Sensitivity and Semantic Minimalism: New Essays on Semantics and Pragmatics,* 112-133. Oxford University Press.

Maitra, I. 2009. Silencing Speech. *Canadian Journal of Philosophy.* 39 (2). 309-338.

Malmgren, A. 2006. Is There A Priori Knowledge by Testimony?. *Philosophical Review* 115 (2). 199-241.

McMyler, B. 2013. The Epistemic Significance of Address. *Synthese.* 190. 1059-1078.

Mercier, H., Sperber, D. 2009. Intuitive and reflective Inferences. From Evens, J. and Frankish, K (Eds), *In Two Minds: Dual Processes and Beyond.* 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.

Moran, R. 2005a. Problems of Sincerity. *Proceedings of the Aristotelian Society.* 105. (3). 341-361

Moran, R. 2005b. Getting Told and Being Believed. *Philosopher's Imprint.* 5. 5: 1-29

Neale, S. 1992. Paul Grice and the Philosophy of Language. *Linguistics and Philosophy.* 15. 509-559

Neale, S. 2005. Pragmatism and Binding. From Szabó, Z (Ed), *Semantics Versus Pragmatics,* 165-285. Oxford University Press

Neta, R. and Rohrbaugh, G. 2004. Luminosity and the Safety of Knowledge. *Pacific Philosophical Quarterly.* 85: 396-406.

Olson, D. Torrance, N. 1987. Language, Literacy, and Mental States. *Discourse Processes.* 10. 157-168.

Onishi, K., Baillargeon, R. 2005. Do 15 month Old Infants Understand False Beliefs? *Science* 308 (5719).

Payne, B. K. 2001. Prejudice and Perception: The Role of Automatic and Controlled Processes in Misperceiving a Weapon. J*ournal of Personality and Social Psychology* 81, No 2. 181-192.

Perrini, E. 2009. Does Contextualism Make Communication a Miracle? *Manuscripto*. 32. (1). 231-247.

Perry, J. 2001. Reference and Reflexivity. Stanford:CSLI Publications.

Pettit, D. 2002. Why Knowledge is Unnecessary for Understanding Language. *Mind*. 111. 443. 519-550.

Pettit, Dean. 2010. On the Psychology and Epistemology of Speech Comprehension. *Baltic Yearbook of Cognition, Logic, and Communication.* 5. 1-43.

Predelli, S. 1998a. 'I am not here now'. *Analysis*. 58. (2). 107-115,

Predelli, S. 1998b. Utterance Interpretation and the Logic of Indexicals. *Mind and Language*. 13. (3). 400-414.

Predelli, S. 2002. Intentions, Indexicals, and Communication. *Analysis*. 62. (4). 310-316.

Predelli, S. 2011. I am Still Not Here Now. *Erkenntnis*. 74. 289-303.

Recanati, F. 2002. Does Linguistic Communication Rest on Inference?. *Mind and Language*. 17. 105-126.

Recanati, F. 2004. Literal Meaning. Cambridge: Cambridge University Press.

Recanati, F. 2010. Truth Conditional Pragmatics. Oxford University Press.

Reimer, M. 1991. Demonstratives, Demonstrations, and Demonstrata. *Philosophical Studies*. 63. 187-202.

Rescorla, M. 2009. Assertion and its Constitutive Norms. *Philosophy and Phenomenological Research*. 79. (1). 98-130.

Richeson, J. A., & Shelton, J. N. 2003. When prejudice does not pay: Effects of interracial contact on executive function. *Psychological Science*. 14 (3). 287-290.

Richeson, J., Trawalter, S., & Shelton, J. 2005. African-Americans' racial attitudes and the depletion of executive function after interracial interactions. *Social Cognition*. 23(4). 336-352.

Richeson, J. A., & Shelton, J. 2007. Negotiating interracial interactions: Costs, consequences, and possibilities. *Current Directions in Psychological Science*. 16(6). 316-320.

Romdenh-Romluc, K. 2002. Now the French are Invading England. *Analysis*. 61. (1). 34-41

Romdenh-Romluc, K. 2006. 'I'. *Philosophical Studies*. 128. 257-283.

Rose, D., Schaffer, J. 2013. Knowledge Entails Dispositional Belief. *Philosophical Studies*. 166. 19-50.

Ross, A. 1986. Why Do We Believe What We Are Told. *Ratio*. 28 (1). 69-88.

Ross, L., Greene, D., House, P. 1977. The False Consensus Effect: An Egocentric Bias in Social Perception and Attribution Processes. *Journal of Experimental Social Psychology*. 13. 279-301.

Rubio-Fernández, P. 2008. On the automaticity of egocentricity: a review of the Egocentric Anchoring and Adjustment model of perspective taking. *UCL Working Papers in Linguistics*. 20. 247-274.

Rysiew, Patrick. 2007. Beyond Words: Communication, Truthfulness, and Understanding. *Episteme*. 4 (3). 285-304.

Rysiew, P. 2008. Rationality Disputes - Psychology and Epistemology. *Philosophy Compass*. 3 (6). 1153- 1176.

Samuels, R., Stitch, S., Bishop, M. 2002. Ending The Rationality Wars: How to Make Disputes About Human Rationality Disappear. From Elio, R (Ed), *Common Sense, Reasoning and Rationality,*236-268. Oxford University Press.

Saul, J. 2012. Just Go Ahead and Lie. *Analysis.* 72 (1). 3-9.

Saul, J. 2012. Lying, Misleading, and What is Said. Oxford University Press.

Schiffer, S. 1992. Belief Ascription. *Journal of Philosophy.* 89 (10): 499-521.

Scott, M. 2010. Religious Language. *Philosophy Compass*. 5 (6). 505-515.

Scott, R., Baillaregon, R. 2013**.** Do Infants Really Expect Agents to Act Efficiently: A Critical Test of the Rationality Principle. *Psychological Science* 24 (4). 466-474.

Sennet, A. 2008. The Binding Argument and Pragmatic Enrichment, or, Why Philosophers Care Even More than Weathermen about 'Raining'. *Philosophy Compass*. 3 (1). 135-157.

Searle, John. 1969. Speech Acts. Cambridge University Press.

Sidelle, A. 1991. The answering machine paradox. *Canadian Journal of Philosophy* 21. 525–539.

Siegel, S. 2012. Cognitive Penetrability and Perceptual Justification. *Nous.* 42 (2). 201-222.

Siegel, S. 2013. The Epistemic Impact of Etiology of Experience. *Philosophical Studies.* 162. 697-722.

Siegel, S. 2013. Can Selection Effects on Experience Effect its Rational Role?. In T. Gendler (Ed.) *Oxford Studies in Epistemology, Vol 4,* 240-270. Oxford University Press.

Soames, Scott. 2012. Philosophy of Language. Princeton University Press.

Sosa, E. 1999. How to Defeat Opposition to Moore. *Philosophical Perspectives.* 13. 141-153.

Sperber, D. and Wilson, D. 1986. Relevance: Communication and Cognition. Oxford: Blackwell; Cambridge, MA: Harvard University Press.

Sperber, D. and Wilson, D. 2004. Relevance Theory. From Horn and Ward (Eds), *The Handbook of Pragmatics,* 607-633. Oxford: Blackwell.

Sperber, D. and Wilson, D. 2002. Pragmatics, Modularity, and Mind Reading. *Mind and Language.* 17. 3-23.

Sperber, D., Clement, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., Wilson, D. 2010. Epistemic Vigilance. *Mind and Language.* 25 (4). 359-393.

Travis, C. 1985. On What is Strictly Speaking True. *Canadian Journal of Philosophy.* 15 (2). 187-229.

Stanley, J. 2000. Context and Logical Form. *Linguistics and Philosophy.* (23). 391-434.

Stanley, J. 2005. Hornsby on the Phenomenology of Speech. *Proceedings of the Aristotelian Society: Supplementry Volumes.* 79. 131-145.

Stanley, J. and Szabó, Z. 2000. On Quantifier Domain Restriction. *Mind and Language.* 15. (2): 219-261.

Savitsky, K, Keysar, B., Epley, N., Carter, T., and Swanson, A. (2010). The Closeness Communication Bias: Increased Egocentrism Among Friends Versus Strangers. *Journal of Experemental Social Psychology.* 47. 269-273.

Stokes, D. 2012. Perceiving and Desiring: A new look at the cognitive penetrability of experience. *Philosophical Studies.* 158. 479-492.

Stokke, A. 2010. Intention-Sensitive Semantics. *Synthese.* 175. 383-404.

Sumner, M., Kim, S., King, E., McGowan, K. (2014). The Socially Weighted Encoding of Spoken Words: A Dual-Route Approach to Speech Perception. *Frontiers in Psychology.* 4. 1-13.

Travis, C. (1985). On What is Strictly Speaking True. *Canadian Journal of Philosophy*. 15 (2). 187-229.

Tversky, A., Kahneman, D. 1974. Judgement Under Uncertainty: Heuristics and Biases. *Science*. 185. 1124-1131.

Van Berkum, A., Van Den Brink, D., Tesink, C., Kos, M., Hagoort, P. 2008. The Neural Integration of Speaker and Message. *Journal of Cognitive Neuroscience*. 20 (4). 580-591.

Van Boven, L., Kruger, J., Savitsky, K., Gilovich, T. (2000). When social worlds collide: overconfidence in the multiple audience problem. *Personality and Social Psychology Bulletin*. 26. 619−628.

Van Fraassen, B. 1980. The Scientific Image. Oxford University Press.

Vorauer, J., & Cameron, J. 2002. So close, and yet so far: does collectivism foster transparency overestimation? J*ournal of Personality and Social Psychology.* 83, 1344−1352.

Wellman, H., Cross, D. Watson, J. 2001. Meta-Analysis of Theory of Mind Development: The Truth about False Belief. *Child Development* 72 (03)*.* 655-684

Wettstein, H. 1981. Demonstrative Reference and Definite Descriptions. *Philosophical Studies* 40 (2). 241-257.

Williamson, T. 1997. "Replies to Commentators." *Philosophy and Phenomenological Research.* 57. 945–53.

Wimmer, H., Perner, J. (1983). Beliefs about beliefs: representation and the containing function of wrong beliefs in young children's understanding of deception. *Cognition*. 13. 103-128.

Wright, C. 2004. On Epistemic Entitlement: Warrant for Nothing (And Foundations for Free?). *Aristotelian Society Supplementary Volume*. 78. 167-212.