

Credibility Deficits, Memory Errors and the Criminal Trial

1: Introduction

Work on testimonial injustice has led to increased recognition that people can suffer a credibility deficit due to aspects of their social identity (see, e.g., Collins 2000; Fricker 2007; Mills 2007; Kidd and Carel 2014). When testimonial injustice occurs people are given less credibility than **they would otherwise** when providing testimony due to the operation of stereotypes about members of their perceived social group. *Credibility deficits* of this type can occur within the criminal justice system, when eyewitnesses, claimants, defendants, or expert witnesses provide testimony in criminal proceedings but their testimony is given less credibility **than it would otherwise** due to the operation of stereotypes about their social group.

Miranda Fricker (2007) draws attention to how credibility deficits can occur within the criminal justice system in one of the primary cases that she uses to illustrate testimonial injustice, that of Tom Robinson in *To Kill a Mockingbird*. Robinson is required to give testimony in his defence at a criminal trial when he is accused of assault. Robinson is a Black man in a racist society and, as such, is not treated as credible when he testifies to his innocence. There are serious negative repercussions for Robinson, as he is found guilty when innocent. In the *To Kill a Mockingbird* example, the credibility deficit that occurs is due to prejudice on the part of the factfinders in the criminal trial. They do not believe the testimony of a Black man due to prejudice against people of his racial group.

Prejudice is a crucial component of testimonial injustice on Miranda Fricker's (2007) extremely influential account. Testimonial injustice only occurs in the presence of social identity prejudice. Prejudice is defined as 'judgements, which may have a positive or negative valence, and which display some (typically epistemically culpable) resistance to counter-evidence owing to some affective investment on the part of the subject' (p. 35). Therefore the credibility deficits occurring within the criminal justice system that are highlighted by Fricker's discussion of testimonial injustice are those involving some resistance to counter-evidence due to an affective investment of the person who is displaying prejudice.

In this paper, I argue that the focus on cases involving prejudice, while incredibly important, has the potential to occlude various cases of credibility deficit that can have a significant detrimental effect on the criminal justice system, leading to poor outcomes within this system for members of negatively stereotyped social groups, and leading important and useful testimony to be wrongly dismissed. My argument highlights previously unrecognised ways that, even in the absence of prejudice, stereotypes can operate to cause the testimony of members of some stereotyped social groups to be wrongly discredited within the context of the criminal trial.¹ These credibility deficits can produce harms that are epistemic, legal and moral. Yet those focusing on credibility deficits that occur due to prejudice are unlikely to give them due attention.

The discussion has the following stages. I first identify a type of credibility deficit that any person can suffer simply due to the way that human memory systems operate. I then show that people who are members of groups that are negatively stereotyped as *cognitively unreliable*, that is, as possessing a cognitive deficiency that makes them unreliable sources of

¹ Fricker recognizes that there can be credibility deficits in the absence of prejudice (2007, p. 21), but denies that these are cases of testimonial injustice, arguing that they are the result of epistemic luck and not injustice. My aim is to identify some cases of credibility deficit that occur in the absence of prejudice that might be obscured by focus on cases that involve prejudice. Why should we think that this might happen? The notion of testimonial injustice has been highly influential but the phenomena discussed in this paper have not been.

knowledge, or *untrustworthy* are especially likely to suffer this type of credibility deficit. We shall see that where people suffer the target credibility deficit due to stereotypes about their groups, there are important similarities to standard cases of testimonial injustice discussed in the literature. However, in a significant subset of cases, the credibility deficits occur in the absence of prejudice, so these cases are not captured by the notion of testimonial injustice as defended by Fricker. Focusing on testimonial injustice as originally conceived by Fricker therefore obscures the importance of the types of cases discussed here. I argue that this calls for one of two types of action to be taken: (a) expansion of the application of the notion of testimonial injustice to include cases of credibility deficit occurring in the absence of prejudice, or, preferably, (b) recognition of a previously unacknowledged type of epistemic, legal and moral harm, sharing many characteristics with testimonial injustice, that occurs in the absence of prejudice.

Before we begin, a note on terminology is required. *Credibility deficit* could be taken to mean a number of different things. It might be thought that a credibility deficit in the technical sense meant by Fricker only occurs when a person is stereotyped and it is as a result of the stereotyping that they are given less credibility than would otherwise. Alternatively, it might be thought that credibility deficits only occur where less credibility is given than is warranted by the evidence available to the person making the credibility judgement. However, Fricker is committed to credibility deficits being possible in the absence of stereotyping and where there is good evidence in support of the credibility assignment that is made. For example, she describes how the fictional character Lieutenant Columbo suffers a credibility deficit when he lures those he is investigating into a false sense of security through his bumbling and shambolic style. A person need not be stereotyping to judge Columbo to be lacking credibility given that he is bumbling and shambolic. The judgement that involves a credibility deficit could be well supported by the evidence that he is bumbling and shambolic. In this chapter the label *credibility deficit* is used in a way consistent with Fricker's definition, to apply to all cases in which a person is treated as less knowledgeable, well-informed and reasonable than they would be if the credibility assignment reflected their true knowledgeability, level of well-informedness and reasonableness. Under this definition, a person can suffer a credibility deficit in a situation in which the person judging them has good evidence in support of the credibility assessment that they make (think again of the Columbo case). This definition of credibility deficit leaves open the possibility that the phenomenon can occur in the absence of prejudice, as a part of a rational response to evidence about the credibility of the person assessed, a point which will become increasingly important later.

2: The Overcritical Juror Argument

Let us begin, then, by identifying the type of credibility deficit that is central to the current discussion. The credibility deficit has been identified in previous work in which the *overcritical juror argument* has been presented and defended (Puddifoot 2020). The overcritical juror argument goes as follows.

Eyewitnesses are prone to making errors in their testimony due to the way that human memory systems ordinarily operate. Human memory systems are fallible in ways that lead people to have false² or distorted memories about past events, which make them susceptible to making errors when sincerely testifying about what they have witnessed. For example, human memory systems are prone to the *misinformation effect*: people form false

² Some might balk at the suggestion that there can be false memories because they are committed to the idea that memories are factive. I believe that there can be memories that are false and distorted but it is beyond the scope of the current discussion to defend this point. For current purposes, I will use the term 'memory' to apply to false and distorted mental representations targeting the past, following common usage in the psychological literature on memory. However, those committed to the factive view of memory can read the same points as referring to some memory-like mental state.

representations of details of events that they have experienced due to being influenced by inaccurate information provided to them after the event by, for instance, other eyewitnesses or suggestive police questioning (see, e.g. Loftus and Palmer 1974; Loftus et al 1978; Okado and Stark 2005). The other eyewitness or police officer might, for example, supply the witness with false information about the speed at which a car was travelling before crashing into another vehicle. The eyewitness might then testify that they saw the car travelling at the speed suggested by the false information supplied by the other person. Due to memory biases like the misinformation effect, eyewitnesses are susceptible to misremembering some details of a crime.

Factfinders are likely to infer on the basis of these errors that eyewitnesses are either *cognitively unreliable*, in the sense that they lack a good supply of information about the event, or *untrustworthy*, in the sense that they are motivated to deceive to ensure that a certain verdict is achieved. Studies show that where an eyewitness provides details that are refuted, reduced credibility is given to their testimony as a whole. This occurs even where the details that are refuted are trivial (Borckardt et al 2003). Under such conditions, jurors are unlikely to convict on the basis of the testimony (Hatvany and Strack 1980; Berman and Cutler 1996), even where the details that were refuted were not crucial to any assessment of the guilt or innocence of the suspect (Berman et al. 1995). Eyewitnesses are assumed, based on a small number of trivial errors, to lack a generally good supply of true memories. What this means is that errors like those made in the misinformation effect, which occur due to the ordinary operation of human memory systems, can lead the whole of the testimony of an eyewitness to be dismissed as lacking credibility.

It is perhaps understandable that factfinders reduce the credibility that they give to the whole of an eyewitness' testimony, choosing not to convict on the basis of the testimony, where some details in the testimony are refuted. However, eyewitnesses nonetheless suffer a credibility deficit when the whole of their testimony is given less credit than it would otherwise due to them providing false details. More specifically, they suffer a credibility deficit when the whole of their testimony is given less credit than it would otherwise due to errors that occur due to them being subject to the *misinformation effect*. Why is this? Because the misinformation effect occurs due to features of human cognitive systems that facilitate being a good eyewitness. Rather than indicating that an eyewitness is unreliable and/or untrustworthy, errors due to the misinformation effect simply indicate that the eyewitness has ordinary human memory systems, found across the human species, which in fact play an important role in the cognitive functioning that makes us good at being eyewitnesses.

The idea that memory errors can be the result of the ordinary operation of cognitive systems that facilitate being a good eyewitness is complex and requires further explanation. The first point to understand is that the misinformation effect, and other similar memory errors, are now generally taken to be due to the *constructive nature of human memory systems* (see e.g. Loftus 2005; Schacter et al. 2011; Michaelian 2013). Memory does not operate like a storehouse, storing discreet and complete files of information about events which can be retrieved at a later point, instead memory systems store traces of information about events that are constructed into a representation of the past event at the point of retrieval. Take, for example, my experience of meeting a friend at the park last autumn. If memory systems worked like a storehouse, there would be a memory file stored somewhere inside me containing complete information about the event. The memory file would contain information about the date, time, and exact location of the meeting; information about which friend I met; information about what we talked about; information about anything we ate or drank; information whether I enjoyed the experience, and so on. In contrast, according to the constructive view of human memory systems, there is no single file of information, the individual pieces of information are distributed as traces of distinct information across the cognitive systems responsible for memory. When I recall the meeting, my cognitive system

constructs a representation of the event, drawing on these traces of information to form a plausible representation of the past.

This process of construction can lead to errors. In the process of construction traces of information about different events can be combined to form false or distorted representations of a past event. For example, when constructing a mental representation that I take to be a memory of meeting in the park with my friend Mo, I might include details from another meeting in a park with my friend Aria. I might seem to recall that Mo brought a football and we had a kick around when in fact it was Aria who brought a football when we met in the same park in the summer. What is found in this example is an error in the process of memory construction.

Under the constructive view of human memory, the misinformation effect is understood as an error in the process of construction. In the misinformation effect traces of information about more than one event are combined at the point of construction to form a single representation. Traces of information about one's personal experience of the event (e.g. a car crash one has witnessed) become combined with traces of false information provided by testimony received after the event (e.g. false information about the speed of the vehicles involved in the crash), to produce a representation of the initial experience that reflects errors contained in the received information. Therefore, the constructive nature of human memory systems is taken to explain the target memory errors that are the result of the misinformation effect.

The second point to see is that the same feature of human memory systems that leads to the misinformation effect—its constructive nature—also brings epistemic benefits (Puddifoot and Bortolotti 2019; Puddifoot 2020). Epistemic benefits are features of a cognitive system that facilitate the achievement of epistemic goals 'including acquiring new true beliefs; retaining and using relevant information; increasing the coherence of a set of beliefs; and gaining understanding' (Puddifoot and Bortolotti 2019). Authors in debates over the constructive nature of human memory systems identify different functions of constructive memory systems, each of which bring epistemic benefits.

For some constructivists, who are defenders of the *mental time travel view*, constructive memory systems perform the function of allowing people to engage in mental time travel, or the projection of the self into the future (Tulving 1985; Suddendorf and Corballis 2007). Similarly, defenders of the *constructive episodic simulation hypothesis* argue that constructive memory systems facilitate the flexible simulation of future events; that is, they allow people to make predictions about novel future events by drawing on a variety of past experiences (Schacter and Addis 2007, see also Schacter et al. 2007, 2011). Simulationists in philosophy claim that remembering simply is simulating a more or less accurate representation of a past event (Shanton and Goldman 2010; Michaelian 2011, 2016a, b). According to each of these views, the constructive cognitive systems underpinning memory are responsible for both recollection of the past and the projection of oneself, through imagination, into the future. As the future rarely exactly resembles the past, it is extremely useful for these cognitive systems to be able to draw from various different past experiences to make predictions about a future event (to time travel into the future or simulate the future). For example, I might want to predict what it will be like meeting Aria in the park in the autumn. To make this prediction, it would be useful to draw on details of the event of meeting Mo in the park, because I met him in the Autumn, but also, of course, important to draw on details of what it was previously like to meet Aria. The constructive nature of human memory systems facilitates this by storing information in the form of traces that can be re-combined in various ways to construct a representation of the future event.

The constructive nature of the cognitive systems underpinning memory thus brings distinctive epistemic benefits. The ability to predict the future allows people to put themselves in a good position to gain further knowledge and understanding about future events and their

consequences. Consider, for instance, a scientist who can predict the outcome of an experiment. She can ensure that she is prepared with appropriate measuring tools, tools for statistical analysis, and so on. If she can predict problems that might occur during the experiment then she can be prepared to solve those problems. She can therefore eventually be in a good position to gain scientific knowledge due to being able to flexibly predict the future.

On an alternative account of the functions of constructive memory systems, Felipe De Brigard (2014) argues that constructive memory systems perform a broader function than simulation of the future: *rather than only simulating the future, the memory system simulate what might have happened and what might happen in the future*. The simulation of hypotheticals of this type brings epistemic benefits. It facilitates the imagination of events that have not yet been experienced. It is therefore a source of knowledge about what events might happen. It also makes tractable the otherwise intractable computational task of considering possible outcomes by focusing attention on outcomes that are likely to happen or to have happened, thus increasing the chance that correct judgements are made about what might happen or have happened.

The third thing to note is that these epistemic benefits brought by humans having constructive memory systems contribute to us being in a generally good position to provide high quality testimony in criminal trials. Consider what is required to provide high quality testimony. Among other things, someone who provides high quality testimony is likely to (i) focus close attention on features of the crime as it occurs, and (ii) identify details of the crime that will be relevant to future criminal investigations. Take the example of a car crash in which one driver is criminally responsible. Someone who is a good eyewitness is likely to have attended closely to the crash as it was occurring and to have identified which features of the crash will be relevant to future investigations (the speed of the cars, the appearance of the cars, any traffic violations committed, etc.). They might even note down information about these features to ensure that they can supply accurate and relevant information about the event at a later time.

Both (i) and (ii) are facilitated by constructive memory systems. One will only be able to attend to relevant features of a crime and to identify which features will be relevant to future investigations if one can successfully predict the future, by engaging in mental time travel, simulating the future, or engaging in general hypothetical thinking. For example, if one cannot predict that two cars are likely to crash when one sees that one of the cars is being driven erratically then one is unlikely to attend to the crash in the ways that would make one a good eyewitness to the crash. One will only be able to accurately predict *novel* crimes, which have not occurred in exactly the same way before, if one can draw flexibly from information about various past experiences that are relevant. For example, one might not have experienced two cars colliding at exactly the speed and location as the crash one is currently witnessing. However, if one can draw on experiences of crashes experienced elsewhere (perhaps in fiction), or at different speeds (perhaps one has been personally involved in a slow collision in the past), combining these experiences to creatively simulate what is likely to happen in the current crash, then one is more likely to predict what will occur in the current scenario. As constructive memory systems facilitate successful prediction of the future, including novel events in the future, they facilitate being a good eyewitness.

We now have reason to think each of the following. First, the misinformation effect is caused by the ordinary operation of constructive human memory systems. Second, the constructive nature of human memory systems brings epistemic benefits, i.e. facilitating the successful prediction of the future, including novel events, through creative thinking that involves selectively imagining only events that are likely to happen. Third, the epistemic benefits associated with constructive memory systems facilitate people being generally good eyewitnesses. The combination of each of these points suggests that a person can make errors due to the misinformation effect, supplying testimony including errors, but nonetheless be a

very good eyewitness. In fact, the errors can be due to the ordinary operation of cognitive systems that otherwise put the eyewitness in a very good position to gather and supply high quality testimony about a crime. Rather than showing that the eyewitness is generally unreliable, errors in testimony that occur due to the misinformation effect simply illustrate that the eyewitness has ordinary constructive human memory systems, which facilitate humans being generally good eyewitnesses.

The constructivist explanations of our target memory errors thus suggest that those testifying in criminal courts can suffer a credibility deficit.³ They can supply testimony that is generally accurate and otherwise of high quality, and which could increase the chance of a correct verdict being made if it were taken seriously, but the testimony can be dismissed due to a small number of errors contained therein. The errors can be wrongly taken to be indicative of the testifier being cognitively unreliable or untrustworthy when they are actually indicative of the ordinary operation of cognitive mechanisms that facilitate the testifier being a good eyewitness.

3: Stereotypes and the Overcritical Juror

Members of any social group could experience the credibility deficit identified in section two. All human beings have constructive memory systems, so any of us could be subject to the misinformation effect. We could make errors due to being supplied with false information about an event we have experienced. These errors might wrongly be taken to be indicative of a general lack of reliability or trustworthiness, leading to a credibility deficit. However, this section illustrates how members of certain social groups, whose social groups are stereotyped as unreliable or untrustworthy, are especially likely to suffer a credibility deficit of this type. Moreover, they are likely to suffer this type of credibility deficit where others might avoid doing so. Support for these claims comes from a body of psychological research studying how stereotypes impact cognition.

Some social groups are stereotyped as lacking the cognitive capacities that are required to provide high quality testimony in a criminal court. The types of cognitive abilities that are associated with providing high quality testimony include the ability to comprehend situations witnessed, the ability to store, retain and retrieve information in memory, and the ability to articulate what is experienced and remembered. Members of social groups like children, people with learning disabilities, and older people are stereotyped as lacking these abilities.⁴ Other social groups are stereotyped as being untrustworthy and likely to deceive, e.g. ethnic minorities, immigrants, and certain religious groups like Muslims.

Psychological research on how stereotypes impact cognition suggest that the testimony of members of these social groups is especially likely to be given less credibility than it deserves when minor errors are present. Stereotypes have been found to distort responses to evidence, such as evidence of errors in a person's testimony, in a number of relevant ways.

³ The focus in previous work has been on eyewitnesses, but others providing testimony could be subject to similar effects, as long as they have personally experienced a crime and been exposed to misleading information after the experience.

⁴ It might be objected that some of these generalisations are not stereotypes because they are legitimate; for example, children at times have poorer abilities to comprehend situations, so it might be thought to be correct to associate children with poor comprehension. However, stereotypes are here being defined neutrally rather than normatively, as mental states that associate perceived members of some social group or group more strongly than others with certain characteristics. This definition applies to cases in which stereotypes are rationally formed and reflect reality, rather than only applying to inaccurate or poorly formed generalisations. This definition of stereotypes is the definition adopted by Fricker (2007) in her work on epistemic injustice, and is commonplace in cognitive science (for further defence of the adoption of this definition of stereotypes see Beeghly 2015 and Puddifoot ms).

The first relevant line of research suggests that stereotypes can operate to ensure that factfinders notice and remember errors in the testimony of those who are members of groups stereotyped as cognitively unreliable or untrustworthy. Experimental participants have been found to notice features that are consistent with stereotypes and to subsequently better remember them if they stereotype than if they do not (e.g. Cohen 1981; Stangor 1988; Signorella and Liben 1984; Bodenhausen 1988; Levinson 2007).⁵ Stereotype-congruent memory effects, where features that are consistent with a stereotype are well remembered, have been found to be especially common under conditions of high cognitive load (Srull 1981; Hastie and Kumar 1979, Hastie 1981), where a task is demanding because there is a large amount of information to process (Hastie and Kumar 1981; Rojahn and Pettigrew 1992; Stangor and McMillan 1992), and under time constraints (Dijksterhuis and van Knippenberg 1995; Rojahn and Pettigrew 1992; Stangor and McMillan 1992). Factfinders are likely to find themselves under these conditions within the context of a criminal trial: they have to process a large amount of information about the crime, suspects, witnesses, and so on. They have to process the information under conditions of high cognitive load and with time constraints imposed by the criminal process. Factfinders are therefore likely to undergo stereotype-congruency effects and remember errors made by members of groups stereotyped as cognitively unreliable and untrustworthy. As the memory of these errors can lead to the conclusion that a testifier is generally unreliable and untrustworthy, stereotypes can operate to produce credibility deficits. Where people who are not members of social groups that are stereotyped as unreliable or untrustworthy might evade the judgement that they are generally unreliable or untrustworthy, because their errors are not noticed or not remembered,⁶ those who are members of such groups can be wrongly judged in this way due to others stereotyping them.

A second line of relevant research suggests that stereotypes can lead ambiguous evidence about an individual to be interpreted in a way that is consistent with a stereotype rather than as it should be interpreted: as ambiguous (Duncan 1976; Sagar and Schofield 1980; Devine 1989; Gawronski, Geschke, and Banse 2003). What this means for the case of the errors in testimony that are being discussed in this paper is that they are likely to be interpreted as being examples of cognitive unreliability or untrustworthiness when they are made by people who are members of social groups stereotyped as unreliable or untrustworthy. The evidence of the errors in testimony is ambiguous: it could be interpreted in various ways, including as simply being evidence of ordinary memory error that is consistent with an eyewitness being generally reliable. However, due to the operation of the stereotype, for example, of older people or immigrants, the errors are likely to be interpreted as evidence of unreliability or untrustworthiness, producing the credibility deficit associated with the overcritical juror effect.

The final relevant line of research suggests that where a stereotype is in operation behaviours can be taken to be expressions of the wider dispositions of a person when they might otherwise be taken to be the result of local features of the situation (Duncan 1976; Sanbonmatsu, Akimoto, and Gibson 1994). The research suggests that where stereotypes are in operation, they can lead behaviours to be explained in terms of the group membership of the individual. Stereotyped individuals can be assumed to have some traits due to their social group membership. These traits are taken to explain their behaviour. How does this research apply to the issue at hand? Where stereotypes are in operation, errors in testimony that are

⁵ For more on how this type of effect can lead to injustices see the discussion in Puddifoot (forthcoming) of mnemonic injustice.

⁶ The same literature suggesting that people's errors are more likely to be remembered if they are fitting with stereotypes about those people's social groups suggests that errors made by people of other social groups, not stereotyped as cognitively unreliable or untrustworthy, are less likely to be remembered. This suggests that there will be cases where a person's errors are remembered when the errors of people from other social groups would be forgotten.

due to the ordinary operation of constructive cognitive mechanisms could be wrongly explained in terms of there being traits that are possessed by the individual—traits such as general unreliability and untrustworthiness—due to their social group membership. For example, errors made by older people might be explained in terms of older people simply being unreliable like that. As a consequence, all or a large part of their testimony might be unduly dismissed.

There are thus various lines of research on the psychology of stereotyping that suggest that members of social groups stereotyped as cognitively unreliable or untrustworthy can suffer the credibility deficit outlined in section two due to the operation of stereotypes of members of their social group. Their errors can be especially likely to be noticed and remembered where the errors of others would not have been; the errors can be taken to provide unambiguous evidence of unreliability or untrustworthiness when they would otherwise be correctly judged to be ambiguous; and the errors can be taken to be explicable in terms of traits (cognitive unreliability or untrustworthiness) of the individual possessed due their social group membership when they would otherwise not be. Within the context of the criminal trial, then, stereotypes can operate to produce credibility deficits like that described in the overcritical juror argument.

4: A practical example

It is worthwhile illustrating the argument from section three through an example. Dennis is an elderly man living in a residential care home due to his physical disabilities. He is cognitively able. He notices that a staff member in the residential home has been acting suspiciously, entering the rooms of other residents when they are asleep and leaving with items that he did not enter with. Over time, items are reported missing from numerous rooms. Enough evidence is gathered for the case to go to trial with the staff member Dennis suspects being tried for theft. Dennis provides a witness statement to the police and then in court, testifying about what he has seen. His testimony is generally reliable. However, he is asked to provide details of the times of the day when he saw the suspect entering the rooms of other residents. He misremembers these details due to having spoken to another resident who has given him false information about the times at which the suspect had been operating. The factfinders in the criminal case are made aware of this error. The factfinders have some awareness of how people can make small memory errors while being generally reliable. However, they also possess a stereotype associating older people with being cognitively unreliable. They are therefore struck by his memory error. When they assess the weight that they should give to his testimony, they remember the error when they would otherwise not do so. They interpret the error, which is ambiguous between being an error due to the ordinary operation of cognitive systems and evidence of wider unreliability, as evidence of the latter. They take the error to be due to a broader set of dispositions towards cognitive unreliability that Dennis possesses due to his age. The factfinders each conclude that Dennis's testimony as a whole should be dismissed due to its general unreliability.

This case has been constructed for the sake of the current discussion, but there is little reason to doubt its plausibility, or that cases of this type are relatively commonplace. Members of groups who are stereotyped as cognitively unreliable are often vulnerable to being victims of crime. They are likely to be giving testimony about crimes they have experienced as well as being stereotyped as cognitively unreliable.⁷ As human beings, members of this type of group

⁷ Given the existence of hate crimes against groups stereotyped as untrustworthy, such as immigrants and Muslims, they are similarly likely to be victims of crime while being stereotyped as untrustworthy. This an important issue that deserves more attention but henceforth my focus is on cases where people are stereotyped as cognitively unreliable and suffer a credibility deficit due to this stereotype because these are most plausibly examples of credibility deficits in the absence of prejudice. I take it that where

are susceptible to making errors due to the misinformation effect. Therefore, they are likely to be given less credibility than they deserve when providing testimony due to the combination of them making errors due to the misinformation effect and being stereotyped.

5: Overcritical Jurors and Testimonial Injustice

We now have good reason for thinking that stereotypes make already vulnerable individuals especially susceptible to suffering the credibility deficit present in the overcritical juror effect. This phenomenon can have a significant impact on the outcome of criminal trials, by leading testimony that could be decisive to be denied the credibility that it deserves. It is therefore important for those seeking to improve the criminal trial that there are the conceptual tools required to understand the phenomenon. One place where we might look for these conceptual tools is in the literature on testimonial injustice. This literature focuses on how members of stigmatized and marginalised groups can suffer credibility deficits due to stereotypes about their social group. However, as we shall see in this section, the target phenomenon does not fit neatly in the category of cases of testimonial injustice: a significant subset of cases in which the phenomenon occurs do not count as cases of testimonial injustice because they do not involve prejudice.

As already noted, testimonial injustice occurs when members of a stigmatized or marginalized group suffer a credibility deficit due to the operation of a stereotype relating to their social group. We have seen that the target cases as described in sections three and four share this feature: a credibility deficit occurs due to the operation of a stereotype relating to the social group of a testifier. In both the standard cases described in the testimonial injustice literature and our target cases people are incorrectly taken not to be knowers due to their social group membership. In addition to this, in both the cases of credibility deficit described in sections three and four and standard cases of testimonial justice, there are epistemic harms to the person stereotyping and epistemic, ethical and practical harms to the person who is stereotyped.

There are epistemic harms to the person stereotyping in each of the cases of credibility deficit because knowledge that a speaker conveys via testimony is not transmitted from the speaker to the hearer, ‘knowledge that would be passed on to a hearer is not received’ (Fricker 2007: 43). In the cases described in sections three and four, important information relevant to the verdict in a trial is not transmitted to the factfinder from the testifier because the factfinder does not recognise the knowledge as such.

There are also epistemic harms to the person stereotyped in standard cases of testimonial injustice. Fricker describes how a speaker can be ‘wronged in her capacity as knower’ (2007: 1), when a credibility deficit occurs due to the operation of stereotypes. The speaker is dishonoured because their sincerity or competence is attacked when, due to their social identity, they are treated as lacking knowledge that they possess (Fricker 2007, p. 46). Fricker describes the wrong as ‘something that can cut deep’ (Fricker 2007, p. 46) because it involves ‘being insulted, undermined or otherwise wronged in one’s capacity as a knower’ (Fricker 2007, p. 46) which is ‘a capacity essential to human value’ (Fricker 2007, p. 46). We shall return (in section 6) to the question of whether those who suffer the credibility deficits described in sections 3 and 4 are wronged by those who misjudge their credibility. It is clear, however, that those who are treated as if they lack knowledge that they could deliver within criminal proceedings due to their social identity can feel insulted or undermined. It could ‘cut deep’ when, for example, an elderly person, person with learning disabilities, immigrant, or Muslim is treated as if they lack knowledge because of their social status and small memory errors that they display. They could quite reasonably feel that their ability to possess

people are stereotyped as untrustworthy, as having a motive to deceive, this is likely to be due to prejudice.

something of fundamental human value—i.e. knowledge—is being questioned. These are significant epistemic harms. For example, when Dennis is denied credibility, and treated as incompetent, due to his age, he is likely to be deeply hurt, wounded and feel disrespected by this.

There are also secondary epistemic harms associated with testimonial injustice (Fricker 2007, pp. 47-8). If a person is aware that their testimony has been undervalued, they can come to doubt their intellectual abilities and therefore lose knowledge that they already possess. Moreover, they can lose their motivation to engage in epistemic projects, and fail to gain new knowledge that they would otherwise have gained through these projects. Where testifiers in a criminal trial are aware that factfinders are not giving due credit to their testimony they can suffer these epistemic harms.

There are further secondary practical harms that are identified with testimonial injustice. For instance people can be denied career opportunities due to their contribution at work not being given the credit it deserves (Fricker 2007). Or, in the example from the testimonial injustice literature with the most intimate link to the current discussion, a defendant like Tom Robinson might be wrongly convicted due to their testimony not being treated as credible (Fricker 2007). These cases involve an injustice with serious practical consequences. In the cases that are the target of the current discussion—cases in which testimony is dismissed in a criminal trial due to evidence of a limited number of memory errors—there are risks of similar injustices and practical harms. People can be denied the justice that would have been afforded by the criminal justice system if their testimony had been taken seriously. And in some cases (think, for example, of a person testifying about their personal experiences of domestic abuse) they will be made vulnerable to further victimisation due to the perpetrator of a crime being free to commit further crime.

What should we conclude on the basis of these observations? One conclusion that we might draw on the basis of this discussion is that the conceptual tools provided within discussions of testimonial injustice can be applied in cases of the sort outlined in sections three and four. It might be thought that we find testimonial injustice in cases where the testimony of members of stereotyped groups is unduly dismissed due to memory errors that would otherwise be forgotten or rightly interpreted as ordinary errors to which we are all susceptible due to the fallibility of human memory systems. The credibility deficits in such cases could be criticised as cases of testimonial injustice. The problem with this conclusion is that the notion of testimonial injustice, as developed within Fricker's framework, does not apply to a significant subset of our target cases.

The relevant cases are those in which prejudice is not responsible for the credibility deficit. Recall that Fricker's testimonial injustice is underwritten by prejudice, and prejudice is defined as involving resistance to counter-evidence 'owing to some affective investment on the part of the subject' (2007, 35). A subset of cases of the type of credibility deficit described in sections three and four occur without prejudice as so defined.

First of all, some of the cases occur in the absence of a resistance to counterevidence. A person may harbour a stereotype of members of a certain social group that is well supported by their experience. For example, they might harbour the stereotype *older people are less cognitively reliable than younger people*, and this might be supported by their experiences of older and younger people. The stereotype might even be said to be in some sense accurate and to reflect the social reality.⁸ The person might apply the stereotype sensitively; being open to

⁸ Again, for a defense of the view that stereotypes can be accurate or inaccurate see Beeghly 2015. For a description of how accurate stereotypes can produce the types of errors made by the factfinder in Dennis' case (e.g. viewing ambiguous evidence as unambiguous) even if they are applying an accurate

the possibility that specific members of the target stereotyped social group are highly reliable, and only applying the stereotype when evidence of unreliability is present. But they might find evidence of unreliability in the form of memory errors that occur due to the operation of constructive nature of human memory systems. Here, they could respond reasonably to the evidence that they have experienced rather than displaying a resistance to counterevidence both when they form the stereotype and when they conclude that the whole of a person's testimony should be dismissed.

Second of all, the types of cases outlined in section three and four can occur in the absence of an affective investment on the part of the hearer. A factfinder does not need to be affectively invested in finding people of a certain group to be cognitively unreliable for them to respond to evidence of memory errors in a way that is consistent with a stereotype of cognitive unreliability.

Take, for instance, a factfinder responding to Dennis's testimony. The factfinder notices and remembers errors made in Dennis's testimony due to the operation of a stereotype associating older people with cognitive unreliability. She then interprets the errors as fitting with Dennis being cognitively unreliable, and explains the errors in terms of the Dennis possessing the trait of cognitive unreliability due to his age. Dennis suffers a credibility deficit because all of his testimony is dismissed. The factfinder suffers an epistemic harm by failing to receive knowledge Dennis conveys. Dennis suffers a primary epistemic harm through the insult, and feelings of being wounded and disrespected that come with being treated as lacking the ability to give knowledge, and a secondary epistemic harm by losing confidence in his own cognitive abilities. Dennis does not succeed in seeing justice done, so suffers a practical harm. Many hallmarks of testimonial injustice are present.

The factfinder does not, however, interpret and explain Dennis's errors in the way that leads to the credibility deficit due to prejudice. The factfinder has no affective investment in judging Dennis in particular, or older people in general, to be unreliable. Meanwhile, the factfinder displays no serious epistemic dysfunction in possessing or applying the stereotype of older people in this case. The stereotype that the factfinder harbours reflects their experiences of some older people being relatively less cognitively reliable than others. The factfinder applies the stereotype in a sensitive way: being aware that specific older people are highly reliable, and only applying the stereotype when evidence is present that seemingly (but falsely) suggests that Dennis is unreliable. The stereotype is applied in Dennis's case because there seems to be evidence of unreliability, i.e. the errors present in Dennis's testimony. In this type of case, a phenomenon that is extremely similar to testimonial injustice occurs but there is a lack of an epistemic dysfunction that occurs due to an affective investment on the part of the hearer, so the concept of testimonial injustice as developed within Fricker's framework does not apply.

Our focus so far has been on cases of credibility deficit in the absence of prejudice that can occur within the context of a criminal trial. But it is worth noting that these cases are representative of a far wider range of cases in which credibility deficits occur in the absence of any resistance to evidence that occurs due to an affective investment on the part of a hearer.

Take the following case from a mental health context. Abdi Sanati and Michalis Kyratsous (2015) present the case report of J.N., a young Ghanaian lady who presents with psychotic phenomena when two weeks post-partum. Included in her symptoms are a general mistrust, persecutory delusions and hallucinations. At first J.N. would only speak to her partner but over time her delusions reduced in intensity, she showed increased insight into her condition,

stereotype see Puddifoot 2017 and Puddifoot ms. For a defense of the comparative view of stereotypes see Puddifoot ms.

and her attitude towards her persecutory ideas began to change. J.N. then started to become distressed about her relationship with her partner, for example, begging him to come to visit her, and saying that he would come back to her eventually. The clinical team determined on the basis of this behaviour that J.N.'s delusional system had developed. J.N. took the team's response to show that they wanted to portray her as 'the crazy one', and her condition worsened. She started to accuse the nurses of having a relationship with her partner and she was judged to have delusional jealousy. Eventually, however, it became clear that J.N.'s partner had left her for another woman, as in their culture a man was allowed to leave their partner if she started behaving in an erratic way. In this case, J.N. suffers a credibility deficit. Her testimony—expressing knowledge that her partner has left her—is judged to be less credible than it really is and this is plausibly due to a stereotype about people with her condition: that they are irrational and/or disconnected from reality.

Sanati and Kyratsous (2015) conclude that J.N.'s case involves testimonial injustice. However, there is good reason to question this analysis. The clinical team who make the judgement that J.N. does not really know that her partner has left her are not responding in a prejudiced way. They are described as distressed by J.N.'s seemingly worsening condition rather than having an affective investment in dismissing what she says. Moreover, they do not display any obvious epistemic dysfunction when they respond to evidence of J.N.'s distressed testimony about her partner's infidelity, applying the generalisation that people with her condition are disconnected from reality, and concluding that her testimony is not reliable (Jefferson ms.). They are making a seemingly reasonable response to the evidence (a) that she is displaying psychotic symptoms associated with delusions, and (b) that the content of her thought that her partner has left her is consistent with delusional jealousy.

What we have found, then, is that there is a subset of cases of credibility deficit due to the operation of stereotypes, which occur both inside and outside of the context of the criminal trial, that are similar in many respects to cases of testimonial injustice—bringing about the same types of epistemic, legal and moral harm—but which occur in the absence of affective investment and serious epistemic dysfunction on the part of the hearer, and therefore are not properly classified as cases of testimonial injustice as it is currently defined in the literature.

6: Outlining a New Form Epistemic Injustice Can Take

What we have identified in section five is a phenomenon that is extremely similar to standard cases of testimonial injustice, leading vulnerable members of society such as older people, children and people with learning disabilities to be susceptible to the harms associated with testimonial injustice due to the application of stereotypes relating to their social identity. However, we have found that the conceptual resources provided within existing discussions of testimonial injustice do not capture the phenomenon.

The problem is even more serious than this, however. It can be especially difficult to identify what is wrong in cases of credibility deficit like those outlined in sections three to five. People seem to be making reasonable responses to the evidence that they have available to them and lack any negative affective investment. As a result, it would be tempting to dismiss the cases as unproblematic, unimportant or impossible to tackle. If people are doing the best that they can do, the latter thought goes, then there can be little good in focusing on the mistakes that they are making. In other words, the very idea that the hearers in the target cases are not committing testimonial injustice could be taken to show that we should not focus efforts on tackling their behaviour. It might be argued that we should focus on cases of testimonial injustice because these are the cases in which hearers are doing something clearly problematic, important, and that can be corrected.

Two types of response are available at this point. It might be argued that the application of the notion of testimonial injustice should be expanded to capture the target cases.⁹ The stipulation that a credibility deficit only counts as a case of testimonial injustice if it is the result of prejudice could be dropped. Alternatively, it could be accepted that there is a form of epistemically, politically and morally harmful credibility deficit that is closely allied to testimonial injustice but importantly distinct because it occurs in the absence of prejudice. Either of these responses would facilitate criticism of cases where there is a credibility deficit that inflicts significant harm in the absence of prejudice. Such cases could either be criticised on the basis that they are cases of testimonial injustice (on the expanded notion of testimonial injustice) or on the basis that they instantiate the closely allied form of credibility deficit.

There is good reason, however, to prefer the introduction of new conceptual resources to capture cases of credibility deficit that occur due to stereotypes but in the absence of prejudice. Rather than labelling these as cases of testimonial injustice, we should acknowledge their distinctiveness, recognising that there is a distinctive form of harmful credibility deficit that has not previously been acknowledged in the literature. It is important to be able to clearly distinguish cases in which people suffer credibility deficits due to prejudice from those in which credibility deficits occur in the absence of prejudice while recognising that both are harmful and worth tackling. The distinction allows us to recognise how there can be cases of *low or no fault credibility deficits*—cases where the hearer giving an inappropriately low level of credibility is not at fault, epistemically or morally, in the same way as hearers perpetrating testimonial injustice. It provides a lens through which people can see how they and other people are potential perpetrators of harmful credibility deficits even where there is seemingly good evidence in support of their credibility judgements and where they lack any affective investment in the judgements of credibility that they make. Put another way, the introduction of new conceptual resources could signal how a hearer can inflict significant epistemic, legal and moral harms like those associated with testimonial injustice while responding to the evidence in a way that could easily be viewed as fair and just rather than as unjustly wronging the speaker.

7. Conclusions

We should be alert to how people can suffer credibility deficits that bring significant epistemic and practical harms due to their membership of stereotyped social groups even in the absence of prejudice on the part of the hearer. The argument in this paper presents an example of how this can happen in the context of the criminal trial. Members of social groups stereotyped as cognitively unreliable or untrustworthy can suffer credibility deficits when providing testimony in a criminal court due to their ordinary memory errors being misinterpreted due to the stereotyping. As the target credibility deficits can occur in the absence of prejudice, the notion of testimonial injustice does not apply to some of the cases, and can in fact obscure their significance. Those who find discussions of epistemic injustice compelling because they highlight harms endured by vulnerable members of society should be concerned that the notion of testimonial injustice has the potential to occlude cases of credibility deficit that occur in the absence of prejudice. In response to this problem, the application of the notion of testimonial injustice could be expanded so that the concept applies to cases of credibility deficit that occur in the absence of stereotyping. However, there are advantages to an alternative option: acknowledging that cases of credibility deficit in the absence of prejudice or obvious injustice are importantly distinct phenomena that produce similarly serious harms.

References

⁹ A precedent for this type of move comes from recent work suggesting that the notion of testimonial injustice should apply to cases of credibility excess as well as credibility deficits (Medina 2011).

- Beeghly, E., 2015. What is a stereotype? What is stereotyping? *Hypatia*. **30**(4), 675-691.
- Berman, G. L. and Cutler, B. L. 1996. Effects of inconsistencies in eyewitness testimony on mock-juror decision making. *Journal of Applied Psychology*. **81**,170–7.
- ^[1]^[2]^[3]Berman, G.L., Narby, D. J. and Cutler, B. L. 1995. Effects of inconsistent eyewitness statements on mock-jurors' evaluations of the eyewitness, perceptions of defendant culpability and verdicts. *Law and Human Behavior*. **19**, 79–88.^[1]^[2]^[3]
- Bodenhausen, G.V., 1988. Stereotypic biases in social decision making and memory: Testing process models of stereotype use. *Journal of personality and social psychology*. **55**(5), 726-737.
- Borckardt, J. J., Sprohge, E. and Nash, M. 2003. Effects of the inclusion and refutation of peripheral details on eyewitness credibility. *Journal of Applied Social Psychology*. **33**, 2187–97.
- Carel, H. and Kidd, I.J., 2014. Epistemic injustice in healthcare: a philosophical analysis. *Medicine, Health Care and Philosophy*. **17**(4), 529-540.
- Cohen, C.E., 1981. Person categories and social perception: Testing some boundaries of the processing effect of prior knowledge. *Journal of Personality and Social Psychology*, **40**(3), 441-452.
- Collins, P.H., 2000. *Black feminist thought: Knowledge, consciousness, and the politics of empowerment*. New York: Routledge.
- De Brigard, F., 2014. Is memory for remembering? Recollection as a form of episodic hypothetical thinking. *Synthese*. **191**(2), 155-185.
- Devine, P.G., 1989. Stereotypes and prejudice: Their automatic and controlled components. *Journal of personality and social psychology*. **56**(1), 5-18.
- Duncan, B.L., 1976. Differential social perception and attribution of intergroup violence: Testing the lower limits of stereotyping of blacks. *Journal of personality and social psychology*. **34**(4), 590-598.
- Fricker, M., 2007. *Epistemic injustice: Power and the ethics of knowing*. Oxford: Oxford University Press.
- Gawronski, B., Geschke, D. and Banse, R., 2003. Implicit bias in impression formation: Associations influence the construal of individuating information. *European Journal of Social Psychology*. **33**(5), 573-589.
- Hastie, R. and Kumar, P.A., 1979. Person memory: Personality traits as organizing principles in memory for behaviors. *Journal of Personality and Social Psychology*. **37**(1), 25-38.
- Hastie, R., 1981. Schematic principles in human memory. *Social cognition: the Ontario symposium*. **1**, 39-88.
- Hatvany, N. and Strack, F. 1980. The Impact of Discredited Key Witness. *Journal of Applied Social Psychology*. **10**, 490–509.^[1]^[2]^[3]
- Levinson, J.D., 2007. Forgotten racial equality: Implicit bias, decision-making, and misremembering. *Duke LJ*. **57**, 345.
- Loftus, E.F. and Palmer, J.C., 1974. Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of verbal learning and verbal behaviour*. **13**(5), 585-589.
- Loftus, E.F., 2005. Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning & memory*. **12**(4), 361-366.

- Loftus, E.F., Miller, D.G. and Burns, H.J., 1978. Semantic integration of verbal information into a visual memory. *Journal of experimental psychology: Human learning and memory*. **4**(1), 19-31.
- Medina, J., 2011. The relevance of credibility excess in a proportional view of epistemic injustice: Differential epistemic authority and the social imaginary. *Social Epistemology*. **25**(1), 15-35.
- Michaelian, K., 2011. Generative memory. *Philosophical psychology*. **24**(3), 323-342.
- Michaelian, K., 2013. The information effect: Constructive memory, testimony, and epistemic luck. *Synthese*. **190**(12), 2429-2456.
- Michaelian, K., 2016. Confabulating, misremembering, relearning: The simulation theory of memory and unsuccessful remembering. *Frontiers in Psychology* [online]. **7** [viewed 1 August 2019]. Available from: doi: 10.3389/fpsyg.2016.01857
- Michaelian, K., 2016. *Mental time travel: Episodic memory and our knowledge of the personal past*. MIT Press.
- Mills, C., 2007. White ignorance. In Sullivan, S. and Tuana, N. eds., *Race and epistemologies of ignorance*. SUNY Press. pp 26-31.
- Okado, Y. and Stark, C.E., 2005. Neural activity during encoding predicts false memories created by misinformation. *Learning & Memory*. **12**(1), 3-11.
- Puddifoot, K. and Bortolotti, L., 2019. Epistemic innocence and the production of false memory beliefs. *Philosophical Studies*. **176**(3), 755-780.
- Puddifoot, K., 2017. Stereotyping: the multifactorial view. *Philosophical Topics*. **45**(1), 137-156.
- Puddifoot, K., 2020. Re-evaluating the credibility of eyewitness testimony: the misinformation effect and the overcritical juror. *Episteme*. **17**(2), 255-279.
- Rojahn, K. and Pettigrew, T.F., 1992. Memory for schema-relevant information: A meta-analytic resolution. *British Journal of Social Psychology*, **31**(2), 81-109.
- Sagar, H.A. and Schofield, J.W., 1980. Racial and behavioral cues in Black and White children's perceptions of ambiguously aggressive acts. *Journal of personality and social psychology*. **39**(4), 590-598.
- Sanati, A. and Kyratsous, M., 2015. Epistemic injustice in assessment of delusions. *Journal of evaluation in clinical practice*, **21**(3), 479-485.
- Sanbonmatsu, D.M., Akimoto, S.A. and Gibson, B.D., 1994. Stereotype-based blocking in social explanation. *Personality and Social Psychology Bulletin*, **20**(1), 71-81.
- Schacter, D.L. and Addis, D.R., 2007. The cognitive neuroscience of constructive memory: remembering the past and imagining the future. *Philosophical Transactions of the Royal Society B: Biological Sciences*. **362**(1481), 773-786.
- Schacter, D.L., Addis, D.R. and Buckner, R.L., 2007. Remembering the past to imagine the future: the prospective brain. *Nature reviews neuroscience*, **8**(9), 657-661.
- Schacter, D.L., Guerin, S.A. and Jacques, P.L.S., 2011. Memory distortion: An adaptive perspective. *Trends in cognitive sciences*, **15**(10), 467-474.
- Shanton, K. and Goldman, A., 2010. Simulation theory. *Wiley Interdisciplinary Reviews: Cognitive Science*. **1**(4), 527-538.
- Signorella, M.L. and Liben, L.S., 1984. Recall and reconstruction of gender-related pictures: Effects of attitude, task difficulty, and age. *Child Development*. **55**(2), 393-405.

Strull, T.K., 1981. Person memory: Some tests of associative storage and retrieval models. *Journal of Experimental Psychology: Human Learning and Memory*, **7**(6), 440-463.

Stangor, C. and McMillan, D., 1992. Memory for expectancy-congruent and expectancy-incongruent information: A review of the social and social developmental literatures. *Psychological Bulletin*. **111**(1), 42-61.

Stangor, C., 1988. Stereotype accessibility and information processing. *Personality and Social Psychology Bulletin*. **14**(4), 694-708.

Suddendorf, T. and Corballis, M.C., 2007. The evolution of foresight: What is mental time travel, and is it unique to humans?. *Behavioral and brain sciences*. **30**(3), 299-313.

Tulving, E., 1985. Memory and consciousness. *Canadian Psychology/Psychologie canadienne*. **26**(1), 1-12.