Social Media, Trust, and the Epistemology of Prejudice

Karen Frost-Arnold

Preprint: Please cite version published in Social Epistemology

Abstract

Ignorance of one's privileges and prejudices is an epistemic problem. While the sources of ignorance of privilege and prejudice are increasingly understood, less clarity exists about how to remedy ignorance. In fact, the various causes of ignorance can seem so powerful, various, and mutually reinforcing that studying the epistemology of ignorance can inspire pessimism about combatting socially constructed ignorance. I argue that this pessimism is unwarranted. The testimony of members of oppressed groups can often help members of privileged groups overcome their ignorance. This paper argues that a particular type of speaker's trust—hopeful trust—can motivate hearers to become cognizant of their privilege and prejudice. I argue that hopeful trust is a powerful way of eliciting trust-responsiveness that can be an effective mechanism for challenging privilege and prejudice. To make this case, I draw on case studies of online attempts to challenge ignorance. While the problems of testimonial injustice, defensive ignorance, and the absence of privilege-cognizant role models can be barriers to challenging ignorance, I show that acts of hopeful trust can effectively overcome these barriers. Thus, hopeful trust can be a useful tool in remedying socially constructed ignorance.

Keywords: Trust, Prejudice, Social Media, Ignorance, Testimonial Injustice

1. Introduction

Ignorance of one's privileges and prejudices is an epistemic problem. As the epistemologies of ignorance literature has grown, it has uncovered a plethora of causes of socially constructed ignorance of privilege and prejudice. Wilful ignorance (Pohlhaus 2012), white ignorance (Mills 2007), testimonial smothering (Dotson 2011), testimonial injustice (Fricker 2007), hermeneutical injustice (Fricker 2007; Pohlhaus 2012), lack of 'world'-travelling (Lugones 2003), and implicit bias (Holroyd 2012) are just some of the causes of this type of ignorance. While the sources of ignorance of privilege and prejudice are increasingly understood, less clarity exists about how to remedy ignorance. In fact, the various causes of ignorance can seem so powerful, various, and mutually reinforcing that reading this philosophical literature can inspire pessimism about our chances of ever remedying these problems. Nonetheless, people do try all the time to educate each other about privilege and prejudice. Many of these attempts are successful, and we need theories that explain their success. Psychologists have been working on this problem.[1] This paper argues that the epistemology of trust and testimony can also make a significant contribution to this work.

I argue that hopeful trust is a powerful way of eliciting trust-responsiveness that can be an effective mechanism for challenging privilege and prejudice. My central question is: how can an oppressed speaker (*S*) effectively use testimony to generate cognizance of privilege or prejudice in a hearer (*H*)? I will show that there are significant epistemic barriers that can prevent *H* from becoming cognizant of privilege or prejudice. The task of the paper is, therefore, to demonstrate how *S*'s testimony can overcome these obstacles. I will argue that *S*'s hopeful trust in *H* can motivate *H* to become aware of their ignorance.

To make this case, I draw on case studies of online attempts to challenge ignorance. The internet is full of blog posts, tweets, Facebook posts, and chat room discussions in which people

discuss privilege and prejudice. People have many reasons for engaging in such discussions. Some want to share their experiences (de Laat 2008) and form community with other activists, while others enjoy being provocateurs, discussing prejudice to get a rise out of their audience. However, much online discussion is not simply self-expression, preaching to the choir, or trolling; instead many people discuss privilege in order to help others unlearn their privileged habits and prejudiced stereotypes (cf. Kahn, Spencer, & Glaser 2013, p.208).

Here are two examples of online attempts to engage in this practice of what I shall call 'challenging privilege and prejudice', by which I mean 'performing speech acts intended to generate cognizance of privilege or prejudice in their audience'. First, in 2012 a Redditor, calling himself european_douchebag, mockingly posted a picture of a college student, Balpreet Kaur, on Reddit. Kaur, due to her Sikh faith, does not trim the hair on her face, and the original Redditor and other Redditors mocked her appearance. When Kaur was alerted to the online taunts, she responded on Reddit with a post explaining her faith and her appearance choices. Kaur's post went viral and was widely praised for the grace with which she handled the online abuse. She wrote the following:

> Hey, guys. This is Balpreet Kaur, the girl from the picture. I actually didn't know about this until one of my friends told on facebook. If the OP [Original Poster] wanted a picture, they could have just asked and I could have smiled :) However, I'm not embarrased or even humiliated by the attention (negative and positive) that this picture is getting because, it's who I am. Yes, I'm a baptized Sikh woman with facial hair. Yes, I realize that my gender is often confused and I look different than most women. However, baptized Sikhs believe in the sacredness of this body - it is a gift that has been given to us by the Divine Being (which is genderless, actually) and, must keep it intact as a submission to the divine will… [B]y not focusing on the physical beauty, I have time to cultivate those inner virtues and hopefully, focus my life on creating change and progress for this world in any way I can. So, to me, my face isn't important but the smile and the happiness that lie behind the face are. :-) So, if anyone sees me at OSU, please come up and say hello. I appreciate all of the comments here, both positive and less positive because I've gotten a better understanding of myself and others from

> this… I hope this explains everything a bit more, and I apologize for causing such confusion and uttering anything that hurt anyone. (as quoted in West 2012)

A second example of challenging privilege and prejudice comes from a blog called *Metamorpho-Sis: A Transgirl's Blog on Philosophy, Feminism, and Transitioning* (http://www.metamorpho-sis.com/blog/). Rachel McKinnon, writing under the pseudonym Philosochick,[2] writes to educate her audience about the prejudices transgender people face. She also writes about cisgender[3] privilege. For example, Philosochick begins her post on difficulties she faces due to having her photo ID not match her new post-transition name with the following: 'This post is just one example of the many situations that transgender people worry about, but cisgender people generally never have to consider' (McKinnon 2012a). Both Balpreet Kaur and Philosochick use social media to try to educate their audience about their privileges and prejudices.

This paper draws on close case studies of these examples to demonstrate the mechanism by which speakers' testimony can challenge privilege and prejudice. In brief, I argue that speakers such as Balpreet Kaur and Philosochick are engaging in hopeful trust. In this hopeful trust, *S* holds out to *H* a vision of an unprejudiced person that *S* trusts *H* to become. Such trust can be effective because *S*'s hopeful trust in *H* can motivate *H* to live up to *S*'s trust. Using these case studies, I show that this hopeful trust is operative in online attempts to challenge privilege and prejudice. However, the implications of this argument extend beyond social media. Hopeful trust can be an effective means of generating cognizance of privilege or prejudice in both online and offline contexts. Demonstrating the efficacy of hopeful trust in online contexts is interesting in its own right, since it illuminates some of the ways that the internet is being used for epistemic and social justice purposes. Additionally, by showing the efficacy of hopeful trust on the

internet, I aim to tackle what, in some ways, is the most difficult test of my hypothesis that hopeful trust has epistemic benefits. Many believe that online trust is impossible or only operative in thick, durable online relationships. But I show that this particular type of trust is both possible on the internet and effective at challenging ignorance in even brief encounters between social media users. Thus, by demonstrating hopeful trust's power in the most difficult test case—brief online interactions—this paper provides a strong argument for the epistemic value of hopeful trust.

The argument proceeds as follows. Section 2 introduces key terminology and delineates the scope of the argument. Section 3 outlines the social epistemology of privilege and prejudice. Three barriers to cognizance of privilege and prejudice are identified: testimonial injustice, defensive ignorance, and the absence of alternatives. This section sets up the problem of how to effectively challenge privilege and prejudice. Sections 4 and 5 present the core of the argument that hopeful trust provides a solution to this problem. Section 4 argues that testimony often involves trust on the part of the speaker. Section 5 argues that a particular type of speaker's trust—hopeful trust—can motivate hearers to become cognizant of privilege and prejudice. Section 6 addresses potential objections to the application of this hopeful trust mechanism to online interactions. Finally, section 7 outlines some of the considerations relevant to determining when and how to use hopeful trust to combat ignorance.

2. Preliminary distinctions and scope of the argument

In this paper, I will use Peggy McIntosh's conception of privilege as 'an invisible package of unearned assets that [the privileged] can count on cashing in each day' (McIntosh

2008, 62). In societies structured by oppression, some groups receive unearned advantages as a result of being at the top of unjust social hierarchies. For example, whites are treated as valued customers, while people of colour are routinely monitored for shoplifting. And cisgender people use bathrooms without worry, while for transgender and genderqueer people, bathrooms are often sites of abuse, humiliation, and even violence. As epistemologists of ignorance have shown, privileged people are often ignorant of their privilege (McIntosh 2008; Mills 2007). The causes of this ignorance are complex. Cultural ideologies promote the myth of meritocracy, allowing privileged people to believe that they have earned their advantages. Moreover, as Charles Mills (2007) has argued, a history of racism in our collective memory promotes versions of history that erase the legacy of oppression that continues to sustain racial inequalities. Similar rewriting of history occurs for other types of oppression.

Prejudice involves holding negative stereotypes of a group: 'A prejudice…is a negatively charged, materially false, stereotype targeting some social group and, derivatively, the individuals that comprise this group' (Begby 2013, 90). In addition to being ignorant of their privileges, people are often also unaware of their prejudices. All of us have likely had experiences with someone who denies they are a racist, while their words and actions speak loudly to the contrary. Additionally, recent research on implicit bias shows that negative prejudicial associations 'can be operative in influencing judgment and behaviour without the conscious awareness of the agent' (Holroyd 2012, 275).

Some qualifications on the scope of my argument are in order. First, prejudice and the habits of privilege have cognitive, affective, conative, and bodily elements (Sullivan 2006, 188). Thus, the process of unlearning one's prejudice or privilege is more complicated than simply changing one's beliefs. As José Medina argues, people's minds, moral characters, and the

structural conditions in which they live all must change for cognizance of privilege and prejudice to occur (Medina 2013, 86). So in claiming that social media can play a helpful role, I am not claiming that simply reading a blog or a post on Reddit will fix one's ignorance. My claim is that it can reduce ignorance. Second, although I argue that oppressed speakers can help the privileged unlearn their ignorance, I am not arguing that the oppressed have an obligation to do so. Third, I discuss privilege and prejudice at a high level of abstraction. I focus on the social-epistemic patterns that many forms of privilege and prejudice share. While I do think there are some such general patterns in the ways that privilege hides itself and our prejudices remain hidden to us, this is not to say that all forms of privilege and all prejudices are the same. So while I will not address the complex differences between the kinds of privileges and prejudices I discuss in the paper, I do want to acknowledge that the differences exist.

3. The social epistemology of privilege and prejudice

In this section, I argue that the social epistemology of privilege and prejudice reveals three *prima facie* barriers to remedying ignorance about privilege or prejudice. These barriers are: (i) the problem of testimonial injustice, (ii) the problem of defensive ignorance, and (iii) the problem of absence of alternatives.

3.1. The problem of testimonial injustice

If hearers are predisposed, by their prejudices and privileges, to distrust the speech of members of an oppressed group, then how can testimony from oppressed people effectively challenge these prejudices and privileges? The problem is that any speech aimed at challenging

ignorance of privilege and prejudice may be discounted due to testimonial injustice. A speaker sustains a testimonial injustice iff 'she receives a credibility deficit owing to identity prejudice in the hearer' (Fricker 2007, 28). When a hearer's stereotypes cause her to grant the speaker's testimony less credibility than would have been granted in the absence of prejudice, the speaker is unjustly undermined in her capacity as a knower (44). Given pervasive problems of testimonial injustice, a social media user from an oppressed group who tries to educate others about the prejudice she experiences may find her speech unduly dismissed. For example, a feminist blogger who describes sexism she experiences in online gaming may be dismissed as exaggerating or whining by other users in the grips of the stereotype of the emotional, irrational woman. Thus, testimonial injustice seems to present a barrier to challenging ignorance. Unrecognized stereotypes may undermine testimony aimed at bringing those stereotypes to the awareness of those in their grips. It is particularly problematic that testimonial injustice undermines the testimony of those most likely to be aware of the harmful effects of the stereotypes (i.e. members of groups targeted by the stereotypes).

3.2. The problem of defensive ignorance

Powerful psychological forces that resist examination of privilege and prejudice present a second problem for attempts to challenge ignorance. Cognizance of privilege involves acknowledging ugly truths about oneself and other members of one's social group(s), and strong psychological forces resist seeing oneself in a negative light. Consider white privilege. Coming to understand white privilege requires coming to grips with a brutal history of slavery, genocide, internment, lynching, segregation, disenfranchisement, theft of land, exploitation, sexual assault,

etc. Whites consciously and unconsciously resist seeing themselves as beneficiaries of such a brutal history (Sullivan 2006).

Even well-intentioned whites, who view themselves as members of a just meritocracy, can fail to see their own role in maintaining oppression. As Shannon Sullivan (2006, 9) puts it, 'the ugliness of a habit can trigger forceful but evasive psychosomatic resistance to conscious examination of it'. For example, Sullivan sees the ideology of colour-blindness as 'an unconscious defensive device that allows white people to avoid recognition of themselves as non-white people often see them… If those who are (allegedly) colorblind have gone beyond race, then it is easy for them to think that racism no longer exists' (127). The existence of such unconscious defence mechanisms can undermine the efficacy of confronting privilege and prejudice.

In fact, while *S* speaks with the aim of generating cognizance of privilege and prejudice, *H*'s response can take several paths contrary to *S*'s aim. First, *H*'s defence mechanisms may cause *H* to simply reject what *S* has to say. Second, in addition to rejecting *S*'s claims, *H* may resent *S* for making *H* uncomfortable, for making *H* feel guilty, or for making what seem to *H* to be unjust and unkind accusations against *H* and her social group. Such resentment may even cause a strengthening of *H*'s prejudice against *S* and *S*'s social group.

3.3. The problem of absence of alternatives

A closely related third problem facing efforts to challenge ignorance of privilege and prejudice is the possibility that the privileged and prejudiced may lack models of alternative unprejudiced or privilege-cognizant identities. Without a positive anti-racist, feminist, anti-

homophobic, class-conscious, or anti-transphobic identity for which she can trade in her ignorance, *H*'s defence mechanisms may kick in. One reason for this is the need to have a positive sense of self and community. Linda Alcoff (2006, 206) alludes to this in the context of whiteness when she says, '"feeling white," when coupled with a repudiation of white privilege, can disable a positive self-image as well as a felt connection to community and history, and generally can disorient identity formation'. This sense of disorientation, due to lack of identity and community, generates fear. For example, civil rights activist Mab Segrest (1994, 80) describes how becoming aware of her whiteness and becoming a self-described race traitor made her afraid of isolation: 'That scared me, too, the possibility of being caught between the worlds of race, white people kicking me out, people of color not letting me in'. In order for whites to sustain cognizance of privilege and to maintain a course of unlearning their prejudice, they need a positive, alternative conception of what it means to be white.

This presents a problem for our social media user, *S*, who aims to generate cognizance of privilege. *S* does not wish to simply preach to the choir; *S* wants to help the privilege-ignorant see their privilege. But it is precisely the privilege-ignorant who are least likely to already have a clear vision of an alternative, privilege-cognizant identity. They are less likely to have role models of the privilege-cognizant and those trying to unlearn their prejudice. And *S* herself cannot be such a role model for *H*, since *S* belongs to the oppressed group.

In sum, there are several significant barriers to challenging ignorance of privilege and prejudice. However, as I argue in the following sections, these barriers can be overcome by acts of hopeful trust on the part of speakers.

4. Testimony and trust

While it has been widely recognized that testimonial exchange requires trust on behalf of the hearer, it has been less recognized that testimony also often depends on speaker's trust in their hearer. To show this, I first outline the kind of trust I have in mind. In testifying to *H*, *S* often, but not always, extends trust to *H*. In the sense under discussion here, trust has two central features. First, trust involves reliance. This makes one vulnerable (Frost-Arnold 2014a; Holton 1994). For example, if I rely on you to keep a secret, then I am vulnerable to damaging disclosures if you divulge my secret. Second, trust involves normative expectations. Trust involves vulnerability to feeling betrayed when the relied-upon party fails to act as expected (Baier 1994, 99). These feelings of betrayal are reactive attitudes that link trust to practices of holding people responsible for their actions (Walker 2006, 80). Thus, when I trust someone to do something, I rely on them to do it, and I do so with normative expectations.

Now, giving testimony often involves speaker's trust, in this sense. Of course, the kind of vulnerability and the types of normative expectations involved depend greatly on the nature of the testimony. Giving testimony can make us vulnerable. We are vulnerable to psychological harm if we are done a testimonial injustice. Similarly, having our speech dismissed or attacked can diminish our standing with others, which can have many other negative consequences both personal and professional. Additionally, often when we tell someone something, we do so with normative expectations. For example, we expect them to listen with an open mind to what we say and take it seriously, we expect them to do what they can to overcome any barriers that might prevent them from hearing us, and we expect them to give us the credibility we deserve. Now consider our examples of testimony that challenges privilege and prejudice. Oppressed speakers risk harm by exposing themselves to the privileged and prejudiced. And they often take this risk with normative expectations that their audience will listen with an open mind and

attempt to overcome whatever barriers may prevent them from giving due credibility. To explain how this type of speaker's trust can motivate change in the hearer, the next section draws on Victoria McGeer's account of hopeful trust.

5. The power of hopeful trust

5.1. What is hopeful trust?

McGeer (2008) provides a compelling account of hopeful trust as a type of trust that can inspire trustworthiness.[4] Hope, she argues, has motivational force. According to psychological hope theory, hope 'is a cognitive and conative activity that involves setting concrete goals, finding pathways to achieve those goals and tapping one's will-power or agency in order to move along pathways to the specified goals' (McGeer 2008, 244). Not only does our hope energize us to achieve our goals, but our hope in others can motivate them. When we engage in hopeful trust, we put ourselves in their hands,

> [a]nd this fact [the fact of putting ourselves in the trustee's hands] is made manifest to them by our very acts and expressions of trust. Hence, by these acts and expressions, we make ourselves vulnerable to them, yes, but in a way that actively holds out a vision to them of what they can be or do. This vision creates for them a kind of affectively charged scaffolding, empowering their own sense of potential agency with the energy of our hope, and thus encouraging them to act in ways commensurate with the vision we maintain. In this way, our hopeful trust can elicit from them an important and powerful kind of trust-responsiveness. (248)

In trusting someone, we hold out to them a vision of the kind of person they can be—a person who lives up to our hopeful vision of caring and competence over the domain of our trust. And this hopeful vision is motivating. The trustee responds to the trust by thinking, 'I want to be as

she sees me to be' (249). In this way, our hopeful vision empowers the trustee to be a kind of role model to herself.

This trust-responsive mechanism is illustrated by an oft-discussed example of what Karen Jones (2004, 5) calls *therapeutic trust*, or 'trust undertaken with the aim of bringing about trustworthiness'. Richard Holton (1994, 63) proposes the following example:

> Suppose you run a small shop. And suppose you discover that the person you have recently employed has just been convicted of petty theft. Should you trust him with the till? It appears that you can really decide whether or not to do so.[5] And again it appears that you can do so without believing that he is trustworthy. Perhaps you think trust is the best way to draw him back into the moral community.

The shopkeeper holds out to the employee a vision of the kind of person he can be. This motivates the employee to live up to this vision, and in so doing the employee is drawn back into the moral community.

The motivational power of the trustor's hopeful vision helps explain the rationality of hopeful trust. Our trust in someone can motivate them to want to be the kind of person who lives up to our trust. By hoping, beyond the other previously available evidence, that the trustee can be this person, we inspire them to become this person. This is what makes hopeful trust rational, when it is rational. While it may be that our evidence, antecedent to our acts of trust, should lead us to believe that the trustee will not act as expected, our very acts of hopeful trust can motivate her to become more trustworthy, and this gives us some evidence that the trustee will live up to our trust. Thus, 'There is nothing rationally inappropriate about extending our trust to others beyond…evidence of their prior trustworthiness, so long as our hopes for what they are capable of in light of our trust are rationally based' (McGeer 2008, 250).

But what makes hope in others rational? According to McGeer (250-1), it depends on 'knowing something about others' values and putative capacities relevant to the domain of our trustful interaction'. This might be just general knowledge about human psychology and knowledge of the ways that hopeful trust can motivate trustworthiness in others (251). In other words, as long as our hopes do not outstrip what it is reasonable to expect people in general to be able to do, and what our hope in them can inspire them to do, then it is rational to hope that they will do it.

We can use the shopkeeper example to summarize the rationality of hopeful trust. The shopkeeper has some evidence, antecedent to any act of trust, that the employee is untrustworthy in this domain (i.e. the employee's past theft). This would seem to make it irrational for the shopkeeper to trust the employee. However, the shopkeeper also knows other things about the employee. The shopkeeper knows that the employee has the common psychological tendencies that spur trust-responsiveness. While trust-responsiveness has just started to receive sustained philosophical attention, it is not an unfamiliar phenomenon to people in general. We are all familiar with two widely manifested psychological characteristics that spur trust-responsiveness. First, people generally do not desire to take advantage of others' vulnerability.[6] Second, we generally want to live up to the expectations of those around us. Now when people trust us, they make themselves vulnerable to us and they expect that we will not take advantage of this vulnerability. Thus, if the shopkeeper has no reason to believe that the employee lacks these psychological features and lacks other reasons that might override the trust-responsive mechanism, then the shopkeeper's hope that the employee will respond to a clear act of trust has some rational basis.

It is important that in this case the hope does not outstrip what it is reasonable to expect people to do. Of course, there are many things that it would be completely unreasonable to hope that someone will do. But if our hopes are that someone will do something that they are able to do and against which there are no overriding incentives, then, in the absence of reasons to believe that the trustee will not respond to our trust, we have at least some reason to believe that our demonstrated trust in them may motivate them to act as we trust them to act.

5.2. Hopeful trust and challenging privilege and prejudice

How does this account of hopeful trust apply to the trust demonstrated by those who seek to challenge privilege and prejudice? The antecedent evidence about the prevalence of testimonial injustice, defensive ignorance, and the absence of alternatives suggests that privileged readers will fail to give due uptake to the claims of such speakers. However, the hopeful trust that the speaker demonstrates in her audience can lead the hearers to become more trustworthy—to work to avoid doing a testimonial injustice, to push past defensive responses, and to seek alternatives. The speaker holds out to the hearer a vision of a kind of person she can be—someone who takes the speaker's words seriously, someone who is aware of her privilege and cognizant of her prejudices. The speaker makes herself vulnerable to the hearer, and this demonstrates hope that the hearer will live up to the speaker's expectations. It is this vulnerability that can trigger the trust-responsive mechanism. When the speaker engages in a clear act of hopeful trust, one that the hearer recognizes as making the speaker vulnerable, the hearer's desire to avoid doing harm can be activated. The hearer desires to avoid doing the

speaker harm and is also motivated to live up to the speaker's vision of the kind of person she can become.

Recall Balpreet Kaur's response to being mocked on Reddit. Like Holton's shopkeeper, Kaur is trusting beyond the antecedent evidence. She has good reason to believe that her audience is closed-minded and intolerant—the community of this subreddit has shown itself to be cruel and prejudiced. But nonetheless she makes herself vulnerable to psychological harm of future attacks. Thus, she engages in hopeful trust.

Now consider what happened in response to Balpreet Kaur's post on Reddit. Not only did the original poster reply, but he apologized:

> I felt the need to apologize to the Sikhs, Balpreet, and anyone else I offended when I posted that picture. Put simply it was stupid. Making fun of people is funny to some but incredibly degrading to the people you're making fun of. It was an incredibly rude, judgmental, and ignorant thing to post.
>
> /r/Funny wasn't the proper place to post this. Maybe /r/racism or /r/douchebagsofreddit or /r/intolerance would have been more appropriate. Reddit shouldn't be about putting people down, but a group of people sending cool, interesting, or funny things. Reddit's been in the news alot lately about a lot of cool things we've done, like a freaking AMA by the president. I'm sorry for being the part of reddit that is intolerant and douchebaggy. This isn't 4chan, or 9gag, or some other stupid website where people post things like I did. It's fucking reddit. Where some pretty amazing stuff has happened.
>
> I've read more about the Sikh faith and it was actually really interesting. It makes a whole lot of sense to work on having a legacy and not worrying about what you look like. I made that post for stupid internet points and I was ignorant.
>
> So reddit I'm sorry for being an asshole and for giving you negative publicity. Balpreet, I'm sorry for being a closed minded individual. You are a much better person than I am Sikhs, I'm sorry for insulting your culture and way of life. Balpreet's faith in what she believes is astounding. (as quoted in West 2012)

Balpreet Kaur's response clearly had an effect on european_douchebag. What strikes me about his apology is how often european_douchebag refers to alternative visions for himself and his

community. He regrets contributing to Reddit's intolerance, and he articulates a vision of Reddit as a place for 'cool, interesting, funny and amazing stuff'. European_douchebag wants to be part of such a community, and it is clear that Kaur's response made him realize that he was not living up to that vision. Notice also that european_douchebag says he has read more about the Sikh faith. Kaur's response gives a thoughtful and personal explanation of her faith, which is a brave and trusting thing to do, given that she is entering a community in which she has been personally mocked. But Kaur does not make a big deal about explaining her faith; she just writes as if her audience will be interested to hear her back-story and will listen with an open mind. And that is what appears to have happened in the case of european_douchebag. He responded to Kaur's vision of her audience as interested and open-minded, and he responded by being interested in the Sikh faith (if we take him at his word). To put it in the language of hopeful trust, Kaur made herself vulnerable by telling her story and challenging the Redditors' prejudice. But she did so in a way that held out to her audience a vision of a community they could be—a space for sharing 'cool and interesting stuff'. And this vision of an open-minded, tolerant community made her audience, in this case european_douchebag, want to live up to that vision.

Having shown how hopeful trust can help educate prejudiced people, I now turn to the question of how hopeful trust may overcome the problems posed by testimonial injustice, defensive ignorance, and the lack of alternatives.

5.3. Avoiding testimonial injustice

When *S* speaks to *H* with the goal of challenging *H*'s privilege or prejudice, *S* often hopes that *H* will take her speech seriously and trusts *H* to do what *H* can to avoid doing *S* a testimonial

injustice. Fricker and others have outlined steps hearers can take to avoid testimonial injustice (e.g. revising upwards one's credibility assessments to more accurately reflect the speaker's trustworthiness) (Fricker 2007; Frost-Arnold 2014a). For those hearers who are aware of the possibility of doing a testimonial injustice, the speaker's hopeful trust that she will be taken seriously can increase their motivation to take these steps.

However, in cases of challenging privilege and prejudice, the hearers are often not even aware that they hold stereotypes that could lead them to discount the speaker's credibility unjustly. In fact, often the speaker's goal is to alert them to the existence of the very stereotypes that could cause them not to take her seriously. Consider a woman testifying to the existence of stereotypes that cast her as prone to overreacting. How can this speech act be successful, given that the audience is likely to be in the grips of the very stereotype that will cause them to discount her testimony? One answer lies in recognizing that ignorance of prejudice is often a result of epistemic laziness and lack of motivation. As Medina puts it, those who grow up with a surrounding social imaginary which includes stereotypes often 'lack the motivation and intellectual curiosity to probe the evidence more fully, to ask about alternative explanations and to find out more. In other words, the social imaginary produces a strong form of epistemic laziness that blocks evidentiary explorations' (Medina 2013, 68). But listening to a speaker who is clearly engaging in hopeful trust can provide additional motivation to probe the evidence, to take a fresh look at one's epistemic habits, and to investigate one's stereotypes. The speaker who, despite some antecedent evidence that she will be dismissed, nonetheless takes the leap of faith in trying to break through to her audience does so out of hope. This hope is empowering and motivating to the hearer, and it can boost the hearer's motivation to probe the evidence of their prejudice.

5.4. Preventing defensive ignorance

Consider one of Philosochick's posts about cisgender people's failure to use the correct names and pronouns for transgender people. Philosochick begins the post with the following, 'This post might make you feel bad, defensive, or angry, and for that I'm sorry' (McKinnon 2012b). She then continues to describe how cisgender privilege often leads cisgender people to develop mental laziness and fail to ensure that they consistently use the new preferred names and pronouns of transitioning and transitioned transgender people. She ends the post with:

> Getting a trans person's name or pronouns wrong *really* hurts: it's like an invalidation of her/his/hir identity. So I implore: a little effort goes a long way. And the good news? The more success you have, the more likely you'll continue to be successful, and the easier it will get.
>
> Thank you in advance. (McKinnon 2012b)

In this blog post we find a speaker who recognizes that her attempt to educate her audience about their privilege runs the risk of making them defensive. Despite this risk, Philosochick seems to believe that her post may still have an effect. She makes a request of her audience that they resist the habits of privilege and instead make an effort to remember transgender people's preferences, and she thanks her audience for what she expects will be their changed behaviour. Thus, I submit that she exhibits hopeful trust in her audience. And I think her trust is effective because her demonstration of hopeful trust can help her audience get past their defensiveness. One reason we respond defensively to challenges to our privilege and prejudices is that few of us like to think of ourselves as bad people. By demonstrating to her reader that she knows they may have a tendency to react defensively, but by thanking them for pushing past that response, Philosochick holds out to her audience a vision of a good person she

believes they can be. This mitigates the sting of the truths she tells her cisgender readers about their learned habits of carelessness, and this lowers the risk of defensive ignorance.

5.5. Providing alternatives

It should be clear how the hopeful trust mechanism offers hearers an alternative vision of the kind of person they can be. This is what the trustor holds out to the trustee when she offers her testimony. The speaker cannot be a role model to the hearer herself; the speaker belongs to the oppressed group, so she cannot be a role model of an anti-racist white or a trans-friendly cisgender person, etc. But the speaker can offer to the hearer a vision of a future version of the hearer, so that the hearer can act as a kind of role model to herself (cf. McGeer 2008, 248).

Additionally, one of the benefits of using social media to challenge prejudice and privilege is being able to invite one's audience into a new community. Many blogs or sub-Reddits form communities of users who interact on a regular basis. By posting on Reddit or allowing comments on one's blog post, speakers can invite their audience to a continued conversation. This can help mitigate one of the features of the problem of absence of alternatives. Recall Mab Segrest's concern that being a white race traitor meant being excluded from both her white community and communities of colour. As Alcoff puts it, becoming cognizant of one's privilege or prejudice can threaten isolation from one's home community. But by offering to be in community with their audience, social media speakers can decrease the threat of isolation.

6. Hopeful trust on the internet

At this point, one might object that I have stretched the concept of hopeful trust into areas where it does not belong. First, one might object to applying the hopeful trust mechanism to internet interactions. Many question whether online trust is ever rational. Primary among the concerns is users' ability to exercise much greater control over their self-presentation online, which can allow for significant deception (McGeer 2004; Nissenbaum 2001). Others argue that trust requires embodied interaction to trigger the affective responses constitutive of full-fledged trust (Dreyfus 2001, 70-1).

In response, I argue that this objection ignores the range of types of online trust. While the absence of embodied interaction and the ease of deception pose problems for those trying to build thick trusting relationships online, they do not pose a serious threat for generalized virtual trust. Much online trust is trust in people in general (Bierhoff & Vornefeld 2004, 51); this is similar to what Trudy Govier calls social trust (Govier 1997). Given what the trustor knows about the average person's dispositions and tendencies, she can rationally place trust in the unknown others with whom she interacts. Consider our trust in the strangers around us on the street not to steal our wallet: 'Given that complete strangers are *complete* strangers, our only basis for expectations about them is some vague and general sense of what "most people" are like, a kind of picture we have of human nature in general' (33). We may walk down the street with some degree of confidence that others will not steal our wallet, because we have a general sense that most people are not thieves. Similarly, some online speakers may have some degree of confidence that most people will not take advantage of a fellow social media user who has made herself vulnerable by sharing something personal in an attempt to educate others. The basis for this expectation is a general sense of what most people are like.

This generalized trust is not completely universal trust in humanity; one can engage in generalized trust with an awareness that not everyone follows the norms of 'most people'. I may walk down the street with confidence that most people are not thieves, but I do take some precautions against the few outliers. Correspondingly, social media speakers can be aware of the diversity of their audience. Some hearers may not have the common psychological tendencies to avoid harming others and to live up to others' expectations; thus, some in the audience may not be susceptible to the trust-responsive mechanism. So a speaker can trust her audience in general to listen with an open mind while still being aware that some will not.

Social media speakers can also recognize another type of diversity within the audience. Some hearers may be so steeped in the ignorance of their privilege and prejudice that even if they do respond positively to the speaker's hopeful trust, the problems of testimonial injustice, defensive ignorance, or absence of alternatives maintain the ignorance. But others in the audience are closer to realizing their ignorance, and these hearers may be pushed to a greater realization of their privilege by a speaker's act of trust. So while the speaker trusts her audience in general, and her hopeful trust may have an effect on many in the audience, it may only be enough of an effect to make a significant difference to some.[7] In sum, I deny that online trust is never possible, and I argue that generalized trust in online strangers is a plausible account of the kind of trust at issue in speech acts that challenge privilege and prejudice.

But this response might prompt a second objection that I am stretching the concept of hopeful trust. One might note that McGeer talks of hopeful trust in the context of thick relationships, such as friendships with a past and an ongoing future. However, I have applied the concept to generalized trust in brief interactions between strangers. While I think this is an innovative extension of the concept, one might think it goes too far because the hopeful trust

mechanism will not work without the kind of thick relationship McGeer analyses. In response, I grant that hopeful trust can be particularly motivating in friendships and other close relationships. Our friends know us well; so if they believe we are capable of change, it is easy to believe that they see something in us that we have missed. Their vision of the person we can be is particularly motivating both because it seems well grounded and because we do not want to let our friends down. Now certainly neither of these factors is as strong in the online relationships I have discussed. However, I do not think they are completely absent. First, a rational trustor is basing her trust on evidence about what people in general are capable of doing. Thus, her vision, as long as it does not appear to be wildly implausible, does appear to the online user as having some grounding. So it appears to the trustee as a vision they could possibly live up to. Second, there can be some motivation to live up to the generalized trust of strangers. When others make themselves vulnerable to us by trusting us, many of us feel loathe (to some degree) to let them down. So while the hopeful trust mechanism may be stronger in thick relationships, I do not think it is wholly absent in thinner ones.

In sum, online acts of hopeful trust that reach a large audience may provide enough motivation to some susceptible audience members that it inspires them to change. It is the possibility for this kind of change that makes it sometimes effective to use social media to challenge privilege and prejudice.

7. Social context and hopeful trust

I have shown that hopeful trust can be an effective mechanism for challenging socially constructed ignorance about privilege and prejudice. This argument, if persuasive, could

motivate more of us to use hopeful trust as a method of combatting ignorance both online and offline. But in using this approach, we face many important questions about when, where, how, and with whom to make ourselves vulnerable. Decisions about when to trust in this way are inherently socially situated. While it is outside the scope of this paper to discuss all of the features of a speaker's social context that shape their decision whether to challenge ignorance with hopeful trust, some of the relevant contours of this decision can be outlined.

One relevant feature of the social context is the speaker's vulnerability. Online harassment is a real and significant threat (Citron 2014; Jeong 2015). Users who share personal stories in acts of hopeful trust make themselves potential targets for trolls,[8] stalkers, and online mobs.[9] Additionally, social media users can experience offline retaliation and persecution for their online speech. Online anonymity (including pseudonymity) can provide critical protection to vulnerable people, including those from marginalized and persecuted groups (Frost-Arnold 2014b). Thus, speakers may choose to share their stories in a spirit of hopeful trust, but they may mitigate some of the risks they face by telling their stories anonymously. It would be interesting to empirically investigate whether such vulnerability reduction is correlated with a reduction in the efficacy of the speech act; in other words, does the trust-responsive mechanism work more when speakers make themselves more vulnerable (and vice versa)? Whether or not it affects the efficacy of the attempt to challenge privilege and prejudice, the choice of using one's offline identity is a pressing decision that speakers from vulnerable groups must face.

Speakers also face the decision of which medium or social media site to use in acts of challenging prejudice and privilege. On sites like Facebook, one can exercise some degree of control over the audience who receives one's posts. In contrast, public blogs or sites like Reddit are, in principle, available to everyone online, even though filtering mechanisms and search

engine prominence may, in practice, reduce one's audience to a smaller subcommunity. Thus, another means of vulnerability reduction for speakers is choice of medium. One might decide to only challenge the ignorance of one's Facebook friends because that is a pre-selected audience whom one already trusts. Of course, the spreadability of digital content means that once one has posted something online, one can never have complete control over who sees it (Grimmelmann 2008). It is always possible that a screenshot or recording of one's act of hopeful trust is shared more widely without one's consent. Nonetheless, speakers' decisions about their audience are shaped by their perceptions of the trustworthiness of the audience.

Decisions about the audience for one's attempts to challenge privilege and prejudice are also shaped by considerations of efficacy and the context sensitivity of generalized trust. Some social media communities are more amenable to acts of hopeful trust than are others. For example, challenging prejudice and privilege on certain parts of Reddit or one's personal blog would often seem much more productive than posting on 4chan's /b/ board, an online home for self-identified trolls who take pleasure in attacking those who they deem 'social justice warriors'. Knowledge of the particular social media community not only shapes one's decisions about the efficacy of challenging ignorance, it also shapes whether one has any generalized trust in the community in the first place. Generalized trust is context-relative. While one may have some generalized trust in people in general, one may not have any trust in members of a particular community. Given what one knows about participants in 4chan's /b/ board or participants in a white supremacist online group, one may have no faith in them to have the kinds of motivations that would make them responsive to one's trust. Thus for some audiences, hopeful trust may be both ineffective and psychologically impossible. These are just some of the considerations relevant to decisions about when to use hopeful trust as a mechanism for challenging privilege

and prejudice. The goal of this paper has been to show that this mechanism can be a powerful tool in combatting socially situated ignorance. It is my hope that future investigations, including empirical studies, will uncover more about when, where, how, and with whom this tool is most effective.

Acknowledgements: I thank Emily Fisher, Rachel McKinnon, and audiences at Hobart & William Smith Colleges, Northwestern University, SUNY Albany, Shimer College, and the 2014 Computer Ethics: Philosophical Enquiry (CEPE) conference.

---

[1] For a recent survey of this literature, see (Lai et al. 2014).

[2] McKinnon asked me to use her real name, rather than her pseudonym, as author of the blog.

[3] 'Cisgender people are those whose gender identity, role, or expression is considered to match their assigned gender by societal standards. Transgender people are individuals who change, cross, or live beyond gender' (Taylor 2010).

[4] While I focus on McGeer's formulation of hopeful trust, several other philosophers have related discussions of trust-responsiveness (Buechner & Tavani 2011; Holton 1994, 63; Jones 2004, 5; Pettit 1995).

[5] Holton's description of the shopkeeper as deciding to trust raises questions about the voluntariness of trust. It is outside the scope of this paper to discuss how choosing to adopt a cognitive attitude of trust is possible. I discuss my account of this phenomenon in (Frost-Arnold 2014a).

[6] Additionally, there is experimental evidence that people are not motivated by self-interest alone (cf. Batson 2002; Gneezy 2005; Mazar, Amir, and Ariely 2008; Ostrom 2005). We are also motivated to avoid doing harm to others. When it comes to dishonesty, there is evidence that 'people care not only about how much they gain from a lie, but also how much the other side loses' (Gneezy 2005, 391).

[7] Generalized trust can also be context-relative; it can be trust in members of a particular community. For example, given my knowledge of my campus community, I might rationally exhibit a generalized trust by leaving my notebook unattended in the library. I trust people in general in this community not to steal it. Similarly, as I discuss in section 7, social media users may choose to engage in acts of challenging privilege and prejudice in some online communities but not others (e.g. posting on Reddit or one's personal blog would seem a much more productive move than posting on 4chan).

[8] For an excellent discussion of the psychology and sociology of trolling, see (Phillips 2015).

[9] For a helpful description of online mobbing, see (Sarkeesian 2012).

References

Alcoff, L. M. 2006. *Visible Identities: Race, Gender, and the Self*. New York: Oxford University Press.

Baier, A. 1994. *Moral Prejudices*. Cambridge, MA: Harvard University Press.

Batson, C. D. 2002. "Addressing the Altruism Question Experimentally." In *Altruism and Altruistic Love: Science, Philosophy, and Religion in Dialogue*, edited by S.G. Post, L.G. Underwood, J.P. Schloss, & W.B. Hurlbut. New York: Oxford University Press.

Begby, E. 2013. "The Epistemology of Prejudice." *Thought: A Journal of Philosophy* 2 (2): 90-99.

Bierhoff, H.-W., and B. Vornefeld. 2004. "The Social Psychology of Trust with Applications in the Internet." *Analyse & Kritik* 26 (1): 48-62.

Buechner, J., and H. T. Tavani. 2011. "Trust and Multi-Agent Systems: Applying the 'Diffuse, Default Model' of Trust to Experiments Involving Artificial Agents." *Ethics & Information Technology* 13 (1): 39-51.

Citron, D. 2014. *Hate Crimes in Cyberspace*. Cambridge, MA: Harvard University Press.

de Laat, P. B. 2008. "Online Diaries: Reflections on Trust, Privacy, and Exhibitionism." *Ethics and Information Technology* 10 (1): 57-69.

Dotson, K. 2011. "Tracking Epistemic Violence, Tracking Practices of Silencing." *Hypatia* 26 (2): 236-257.

Dreyfus, H. L. 2001. *On the Internet*. New York: Routledge.

Fricker, M. 2007. *Epistemic Injustice: Power and Ethics in Knowing*. New York: Oxford University Press.

Frost-Arnold, K. 2014a. "The Cognitive Attitude of Rational Trust." *Synthese* 191 (9): 1957-1974.

Frost-Arnold, K. 2014b. "Trustworthiness and Truth: The Epistemic Pitfalls of Internet Accountability." *Episteme* 11 (1): 63-81.

Gneezy, U. 2005. "Deception: The Role of Consequences." *American Economic Review* 95 (1): 384-394.

Govier, T. 1997. *Social Trust and Human Communities*. Buffalo: McGill-Queen's University Press.

Grimmelmann, J. 2008. "Accidental Privacy Spills." *Journal of Internet Law* Jul.: 8-35.

Holroyd, J. 2012. "Responsibility for Implicit Bias" *Journal of Social Philosophy* 43 (3): 274-306.

Holton, R. 1994. "Deciding to Trust, Coming to Believe." *Australasian Journal of Philosophy* 72 (1): 63-76.

Jeong, S. 2015. *The Internet of Garbage*. Forbes.

Jones, K. 2004. "Trust and Terror." In *Moral Psychology: Feminist Ethics and Social Theory*, edited by P. DesAutels & M. Urban Walker. New York: Rowman & Littlefield.

Kahn, K. B., Spencer, K., and J. Glaser. 2013. "Online Prejudice and Discrimination: From Dating to Hating." In *The Social Net: Understanding Human Behavior in Cyberspace*, edited by Y. Amichai-Hamburger. New York: Oxford University Press.

Lai, C. K., Marini, M., Lehr, S. A., Cerruti, C., Shin, J. E. L., Joy-Gaba, J., Ho, A. K. et al. 2014. "Reducing Implicit Racial Preferences: I. A Comparative Investigation of 17 Interventions." *Journal of Experimental Psychology: General* 143 (4): 1765-1785.

Lugones, M. 2003. *Pilgrimages/Peregrinajes: Theorizing Coalition against Multiple Oppressions*. New York: Rowman & Littlefield.

Mazar, N., Amir, O. and D. Ariely. 2008. “The Dishonesty of Honest People: A Theory of Self-Concept Maintenance.” *Journal of Marketing Research* 45 (6): 633-644.

McGeer, V. 2004. “Developing Trust on the Internet.” *Analyse & Kritik* 26 (1): 91-107.

McGeer, V. 2008. “Trust, Hope and Empowerment.” *Australasian Journal of Philosophy* 86 (2): 1-18.

McIntosh, P. 2008. “White Privilege and Male Privilege.” In *The Feminist Philosophy Reader*, edited by A. Bailey & C. Cuomo. New York: McGraw Hill.

McKinnon, R. 2012a. “Getting my A/V Key for the Semester.” *Metamorpho-Sis: A Transgirl’s Blog on Philosophy, Feminism, and Transitioning*. Accessed 10 May, 2016. http://metamorpho-sis.net/2012/05/3-getting-my-av-key-for-the-semester/

McKinnon, R. 2012b. “Being on Autopilot.” *Metamorpho-Sis: A Transgirl’s Blog on Philosophy, Feminism, and Transitioning*. Accessed 10 May, 2016. http://metamorpho-sis.net/2012/06/5-being-on-autopilot/

Medina, J. 2013. *The Epistemology of Resistance*. New York: Oxford University Press.

Mills, C. 2007. “White Ignorance.” In *Race and Epistemologies of Ignorance*, edited by S. Sullivan & N. Tuana. Albany, NY: SUNY Press.

Nissenbaum, H. 2001. “Securing Trust Online: Wisdom or Oxymoron.” *Boston University Law Review* 81 (3): 635-664.

Ostrom, E. 2005. “Policies that Crowd out Reciprocity and Collective Action.” In *Moral Sentiments and Material Interests*, edited by H. Gintis, S. Bowles, R. Boyd & E. Fehr. Cambridge, MA: MIT Press.

Pettit, P. 1995. “The Cunning of Trust.” *Philosophy & Public Affairs* 24 (3): 202-225.

Phillips, W. 2015. *This is Why We Can’t Have Nice Things: Mapping the Relationship between Online Trolling and Mainstream Culture*. Cambridge, MA: MIT Press.

Pohlhaus, G. 2012. “Relational Knowing and Epistemic Injustice: Toward a Theory of Willful Hermeneutical Ignorance.” *Hypatia* 27 (4): 715-735.

Sarkeesian, A. 2012. “TEDxWomen Talk about Online Harassment & Cyber Mobs.” *Feminist Frequency*, Accessed 3 March, 2016. http://feministfrequency.com/2012/12/05/tedxwomen-talk-on-sexist-harassment-cyber-mobs/

Segrest, M. 1994. *Memoir of a Race Traitor*. Boston: South End Press.

Sullivan, S. 2006. *Revealing Whiteness: The Unconscious Habits of Racial Privilege.* Bloomington, IN: Indiana University Press.

Taylor, E. 2010. “Cisgender Privilege: On the Privileges of Performing Normative Gender.” In *Gender Outlaws: The Next Generation*, edited by K. Bornstein & S. B. Bergman. Berkeley: Seal Press.

Walker, M. U. 2006. *Moral Repair: Reconstructing Moral Relations after Wrongdoing*. New York: Cambridge University Press.

West, L. 2012. “Reddit Users Attempt to Shame Sikh Woman, Get Righteously Schooled.” *Jezebel*, Accessed 23 September, 2012. http://jezebel.com/5946643/reddit-users-attempt-to-shame-sikh-woman-get-righteously-schooled