episteme

ARTICLE

Credibility excess as an epistemic injustice

Keith Dyck

Philosophy, UC Santa Barbara, Santa Barbara, CA, USA
Email: keithdyck@ucsb.edu

Abstract

According to Fricker's (2007) seminal account, an epistemic injustice is done when, based on prejudice, a hearer ascribes to a speaker a level of credibility below what they deserve. When prejudice results in credibility excess, however, Fricker contends no similar injustice takes place. In this paper, I will challenge the second of these claims. Using a modified version of Zollman's (2007) two-armed bandit model, I will show how the systematic over-ascription of credibility within a dominant group can produce epistemic advantages for that group relative to non-group members.

1. Introduction

As the film "The Talented Mr. Ripley" reaches its climax, Marge Sherwood confronts Herbert Greenleaf with her belief that, rather than dying by suicide, his son was murdered by Tom Ripley. Greenleaf responds by rejecting Sherwood's claim, retorting, "Marge, there's female intuition, and then there are facts" (Minghella and Highsmith 2000, 130). According to Fricker's (2007, 9) influential interpretation, Greenleaf's dismissal of Sherwood's testimony provides a clear example of an epistemic injustice. Based on the stereotype that women are prone to make claims based on emotion rather than reason, Greenleaf assigns to Sherwood a level of credibility below what she deserves. In doing so, Sherwood is "ingenuously downgraded and/or disadvantaged in respect of their status as an epistemic subject" (Fricker 2017, 53).

Now, consider a closely related example where, rather than Sherwood, it is Freddie Miles who presents his suspicions regarding the death of Greenleaf's son. If, based on stereotypes associated with being male, Greenleaf immediately accepts Miles' contention as almost certainly true, it would seem that Greenleaf has assigned to Miles a level of credibility in excess of what he deserves.

Many of us feel an intuitive pull towards reaching the same verdict in each of these cases: if the first is an example of an epistemic injustice, surely the second is as well.[1] Fricker (2007) takes this intuition to be based on the tacit assumption that a distributive

---

[1] In Fricker's (1998, p. 170) earliest discussion of epistemic injustice, she affirms this intuition – writing that when, due to prejudice regarding social identity, there is "a mismatch between rational authority and credibility . . . we should acknowledge that there is a phenomenon of *epistemic injustice*." She states in her more developed work, *Epistemic Injustice,* that she has since "changed my mind" (Fricker 2007, 19n14).

notion of justice is what should be operative in making such evaluations. However, according to Fricker (2007, 19), such a notion is only appropriate when what is involved is a good, such as healthcare or wealth, that is "finite and at least potentially in short supply." Since credibility is "not generally finite in this way ... there is no analogous competitive demand to invite the distributive treatment" (Fricker 2007, 20). Greenleaf's over-ascription of credibility to Miles in no way constrains the credibility he can assign to others, so it should be taken to involve no epistemic injustice at all.

The lack of symmetry in Fricker's treatment of the prejudicial misassignment of credibility has recently come under some scrutiny. Medina (2011; 2013), for instance, takes it that while credibility is not a scarce good, judgments of credibility are always comparative in nature. Assessment of injustice is then a matter of proportionality, with credibility excess inextricably linked to credibility deficit and deficit linked to excess. Coady (2017, 62), meanwhile, takes it as "clearly wrong" that credibility is a non-finite good, claiming it would be both irrational and psychologically impossible to assign maximal credibility to every individual in every instance. In fact, he goes on to suggest, credibility is often in short supply and there is frequently competition over its assignment.

In this paper, I will take a different tack in arguing that credibility excess can, at least in some instances, involve an epistemic injustice. While credibility may arguably be a good that is never in short supply, high credence in true propositions is a seemingly more basic epistemic good that often is. Fairness in the distribution of such credence values should then be of concern in assessing the potential injustice involved in the prejudicial over-ascription of credibility.[2] Using a modified version of Zollman's (2007) two-armed bandit model, I will examine the evolution of credence in an epistemic community investigating a matter of fact when credibility is over-ascribed within a dominant group. What simulation results will show is that while such an over-ascription negatively impacts the accuracy of all community members in reaching high credence in the relevant true proposition, the related benefit of reaching high credence earlier in successful learning periods is disproportionately realized by members of the dominant group. Since epistemic agents in the marginalized group are systematically disadvantaged in such cases, the prejudicial assignment of excess credibility should be taken to involve an epistemic injustice.

The structure of this paper is as follows. Section 2 presents Zollman's two-armed bandit model, along with modifications made to the model in order to introduce the potential for credibility excess. Section 3 will then use simulation results produced using this model to show how the over-ascription of credibility within a dominant group can produce relative epistemic advantages for members of that group over non-group members. The paper concludes with some lessons learned and potential areas for further research.

2. Modeling framework

2.1. Zollman's two-armed Bandit model

In examining how credibility excess may influence epistemic outcomes, I will rely on modifications to the two-armed bandit model first introduced into the philosophy of

---

[2]Coady (2010, p. 112) has similarly suggested a form of epistemic injustice involving "injustice in the distribution of the epistemic good of knowledge." Fricker (2017) has recently become more open to suggestions of this type, allowing that epistemic goods such as education, information, or expert advice may be unjustly distributed. See Irzik and Kurtulmus (2021) for a related discussion concerning distributive epistemic injustice in the production of scientific knowledge.

science by Zollman ([2007](#)).[3] As motivation for his model, Zollman considers a community of medical researchers who, at regular intervals, must choose between a well-understood method and a newly introduced method for treating their patients. Since the researchers feel obliged to maximize their patients' chances for a favorable outcome, they always treat their patients using the method they currently believe is best. This assessment, however, may change over time as researchers update their credence as to the relative superiority of the new method of treatment based on reported treatment outcomes.

In modeling this scenario, Zollman uses a network of N agents who, at regular intervals, decide to perform one of two possible actions. By stipulation, action A has a success rate of 0.5 and action B (the "B"etter of the two actions) has a success rate of $0.5 + \varepsilon$. While agents know action A's success rate with complete certainty, they are only certain that action B's success rate is either ε better or worse than that of action A.[4]

When deciding which action to perform, agents opt for the action they currently believe is more likely to be successful – this assessment is based on their credence that the success rate of action B is superior to that of action A. Epistemically, however, agents are only better off if action B is performed, as it is only then that evidence can be gathered as to its relative rate of success.

In simulating how such a community epistemically evolves over time, each agent is initialized with a randomly selected credence as to the superiority of action B in the interval $(0, 1)$. Deciding which action to perform based on this credence, a simulation round commences with each agent performing their preferred action n times. Agents then report the observed successes and failures of their actions, sharing these results with their immediate network neighbors. Using the standard Bayesian method in conjunction with the reported outcomes they have gathered, agents update their credence concerning the superiority of action B. The round ends with agents once again selecting their preferred action based on their updated credence.[5]

Given enough simulation rounds, only two stable outcomes are possible for fully connected networks of agents. If all agents choose to perform action A in a given round, fully stable incorrect consensus results. Since, in such a scenario, new evidence regarding action B will never be produced going forward, credence values for all agents will remain fixed below 0.5. Correct consensus, the only other possible stable outcome, is taken to have been achieved when all agents have credence values above some threshold – with this threshold set at 0.99 for the simulations discussed in this paper.[6] Correct consensus is only approximately stable in that an arbitrarily long sequence of unlucky results when performing action B could, at any point, result in every agent switching to action A. This possibility, however, becomes increasingly unlikely the higher each agent's credence becomes. Polarized outcomes, where agents stably differ in their preferred choice of action, are precluded in this setup.

Running simulations with agents placed in several standard network configurations (e.g. wheel, cycle, complete), Zollman ([2007](#)) uses his model to investigate how changes in how widely reported action results are shared impacts the success of an epistemic community. He

---

[3]Zollman's model is itself a slight modification of the model presented by Bala and Goyal ([1998](#)). A number of philosophers of science, including Wu ([2023](#)); Holman and Bruner ([2017](#)); Weatherall and O'Connor ([2021](#)); O'Connor and Weatherall ([2019](#), [2018](#)), have used Zollman's model, as well as the closely related Zollman ([2010](#)) model, to investigate issues related to epistemic inquiry in a social setting.

[4]Agents also know the value of ε with complete certainty, with ε serving as an adjustable model parameter in the interval $(0, 0.5]$ used to specify how clearly differentiated the two actions are.

[5]Only reported results concerning action B will be involved in calculating an agent's updated credence, $C_{updated}$. Using the standard Bayesian approach, if action B is performed n times with s successes, an agent's initial credence, $C_{initial}$, will be updated as follows: $C_{updated} = \frac{C_{initial}}{C_{initial} + (1 - C_{initial})\frac{0.5 - \phi \cdot 2 \cdot s - n}{0.5 + \phi}}$.

[6]In Zollman ([2007](#)), 0.9999 was used for this threshold instead.

finds that, for at least some range of parameters, the more connected the agents, the less frequently, but faster, (approximately) stable correct consensus is reached.[7]

2.2. Incorporating reliability and credibility

Zollman's motivating example involves medical researchers whose choice between treatment options is influenced by past reported outcomes. Determining a treatment's success or failure, however, is often a less than straightforward task. A patient's condition can improve for reasons unrelated to treatment, medical tests can produce misleading results, researchers can be more or less skilled at interpreting symptoms, and so on. Reported outcomes are then often less than fully accurate, with levels of accuracy varying depending on disease, treatment, and researcher-specific factors.

Incorporating this real-world feature into Zollman's modeling framework, we will take the reliability, r, of an agent to be the probability that an action's success or failure will be reported by the agent as such (with $1 - r$ providing the probability that an agent will misreport an action's outcome as its opposite). Recall that the objective rate of success of action B is $0.5 + \varepsilon$. When action B is performed n times, successes will then follow a binomial distribution with parameters n and $0.5 + \varepsilon$. The successes reported by an agent with reliability r will follow a binomial distribution with parameters n and $0.5 + (2r - 1)\varepsilon$ instead.[8]

Figure 1 provides a concrete illustration of the distributions involved when $n = 150$, $\varepsilon = 0.3$, and $r = 0.6$. The dark gray graph shows the "Actual" distribution of successes when action B is performed 150 times, while the light gray graph shows the distribution of successes "Reported" by an agent with reliability 0.6.
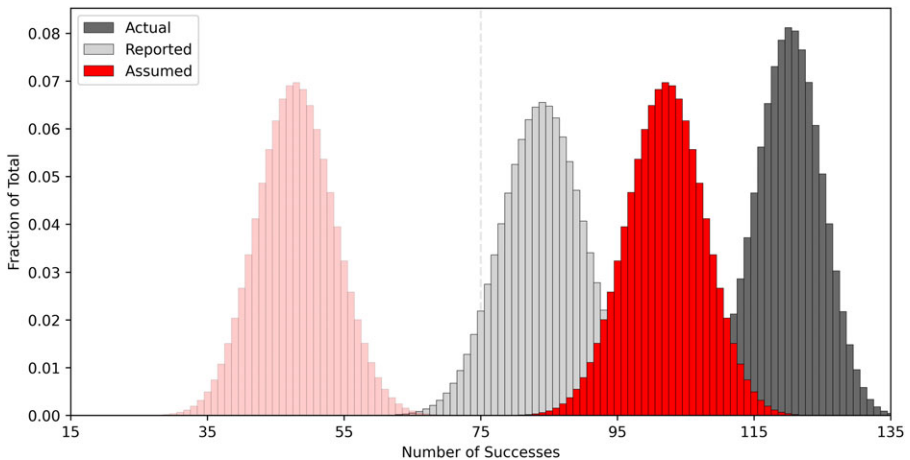

Figure 1. Distributions of actual successes, reported successes, and successes assumed by an updating agent, when $n = 150$, $\varepsilon = 0.3$, $r = 0.6$, and $c = 0.8$.

---

[7]Zollman (2010); Kummerfeld and Zollman (2016); Grim et al. (2013) reach similar conclusions based on computer simulations. Rosenstock et al. (2017); Frey and Šešelja (2020) challenge the robustness of these claims.

[8]The rate at which an agent with reliability r will report successes when performing action B is given by: $(0.5 + \varepsilon) \cdot r + (0.5 - \varepsilon) \cdot (1 - r) = 0.5 + (2r - 1) \cdot \varepsilon$. As expected, this rate matches the objective rate of success of action B when $r = 1$.

Given potential inaccuracies in reported outcomes, it makes sense for an agent to factor in a reporting agent's reliability when updating on the information they provide. Take credibility, c, to be this assumed reliability. When the credibility assigned by an updating agent to a reporting agent matches the reporting agent's actual reliability (i.e. $c = r$), credibility can be said to have been "correctly" ascribed; when $c < r$, credibility has been under-ascribed; and when $c > r$, credibility has been over-ascribed.[9]

Under the assumption that action B is superior, an updating agent who assigns credibility c to an agent performing action B will take reported successes to follow a binomial distribution with parameters n and $0.5 + (2c - 1)\varepsilon$. The dark red graph in Figure 1 shows this "Assumed" distribution when $c = 0.8$. Since, in this case, credibility is over-ascribed (i.e. $c > r$), but not maximally so (i.e. $c < 1$), the bulk of the "Assumed" distribution falls between the bulk of the dark gray "Actual" and the light gray "Reported" distributions. Under the assumption that action B is inferior (i.e. the objective success rate of action B is $0.5 - \varepsilon$), the updating agent will take the light red graph to give the distribution of reported successes instead. This second distribution, also needed when performing Bayesian updating, is simply the "Assumed" distribution horizontally reflected about its center (i.e. about $x = \frac{n}{2} = 75$).

2.3. Over-ascribing credibility within a dominant group

With the Zollman model suitably modified to allow for less than maximally high levels of both reliability and credibility, an examination of the systematic over-ascription of credibility by a group of agents now becomes possible.

In modeling such a scenario, agents will be divided into two mutually exclusive groups, the "dominant" group and the "marginalized" group, with the parameter f used to specify the fraction of agents in the marginalized group. For simplicity in interpreting results, agents will be arranged in a complete network (i.e. all outcome reports will be shared with the entire community), and every agent will be assigned the same reliability, r. When both the updating and reporting agents are members of the dominant group, the credibility ascribed by the updating agent to the reporting agent will be set to $c_{dominant}$. In all other cases, credibility will be set to the reporting agent's actual reliability (i.e. r). The over-ascription of credibility within the dominant group can then be simulated by specifying a $c_{dominant}$ value in the interval $(r, 1]$.[10]

A concern at this point may be whether it is realistic to assume that it is only members of the dominant group that over-ascribe credibility. Should members of the marginalized group be taken to over-ascribe credibility as well? Using data from the well-known Race Attitude Implicit Association Test, Morehouse and Banaji (2024) report that while White Americans on average exhibit a clear pro-White bias, Black Americans on average display very little implicit racial bias.[11] More generally, Morehouse and Banaji (2024, 31) conclude that "unlike members of socially advantaged groups, who consistently display implicit in-group preferences, members of socially disadvantaged groups typically do not." Taking within-group over-ascription of credibility to be a reflection of in-group preferences, the modeling assumption that it is only members of the dominant group who over-ascribe credibility seems like a plausible one.

---

[9]In Zollman's original model $c = r = 1$, so all ascriptions of credibility are correct.

[10]For simplicity in modeling, we will take members of the marginalized group to self-ascribe credibility r and members of the dominant group to self-ascribe credibility $c_{dominant}$ when updating.

[11]Interested readers can take the Race Attitude Implicit Association Test, as well as a variety of other implicit association tests, at https://implicit.harvard.edu/implicit/selectatest.html.

When an updating agent over-ascribes, rather than correctly ascribes, credibility to a reporting agent, credence updates are amplified. Provided r is greater than 0.5 (i.e. the reporting agent is more likely to accurately report an action result than misreport it), the over-ascribing agent will update in the same direction as a similarly epistemically positioned correctly ascribing agent, but with a magnitude that is greater.[12] In this sense, members of the dominant group can be said to be "overly aggressive" or "over-react" when updating on outcomes reported by dominant group members.

3. Simulating the over-ascription of credibility

In producing the results discussed in this section, one hundred thousand simulations were run for every combination of the following parameters:

- Reliability of all agents (r): 0.85
- Credibility assigned within the dominant group ($c_{dominant}$): 0.85, 0.9, 0.95, 1
- Number of agents (N): 12, 24, 48
- Fraction of agents in the marginalized group (f): $\frac{1}{12}, \frac{1}{6}, \frac{1}{3}, \frac{1}{2}$
- Probability of action B's success ($0.5 + \varepsilon$): 0.51, 0.52, 0.55, 0.6, 0.7
- Number of times preferred action is performed per round (n): 1, 5, 10, 20

Simulations where $c_{dominant} = r$ provide a "baseline" for a particular combination of N, f, ε, and n values. Since, in these simulations, all agents correctly ascribe credibility to all agents, no systematic differences in agent-specific outcomes are possible. The epistemic impact of the over-ascription of credibility within the dominant group can then be examined by comparing results from simulations where $c_{dominant} > r$ to the corresponding baseline.

3.1. Accuracy and speed in reaching correct consensus

We begin by examining how the over-ascription of credibility within the dominant group impacts both the frequency of correct consensus and the speed with which correct consensus is achieved. In order to eliminate noisy comparisons involving success rates near the upper limit of 1, only N, f, ε, and n values where the frequency of correct consensus in baseline simulations was below 0.99 will be considered in this subsection.

Figure 2 shows typical results for the fraction of simulations that reach correct consensus as $C_{dominant}$ is varied, with colored lines corresponding to different fractions of agents in the marginalized group. As can be seen, regardless of the marginalized group's relative size, the fraction of simulations ending in correct consensus decreases as $c_{dominant}$ increases from the baseline value of 0.85 to the maximally over-ascribed value of 1. This effect is more pronounced the smaller the relative size of the marginalized group (i.e. the lower the value of f). Both of these trends hold for all parameter values considered.

This decrease in accuracy is accompanied by an increase in the speed with which epistemic success is achieved. For all parameter values considered, the entire community reaches correct consensus in less rounds on average when the dominant group over-ascribes credibility than in the baseline scenario. This effect is more pronounced as the

---

[12]Briefly, this can be shown by taking the updating function provided in footnote 2.1 and replacing ε with $(2c - 1)\varepsilon$. By differentiating the resulting function with respect to c, it can be shown that, when $c > 0.5$, updating magnitudes increase with c.
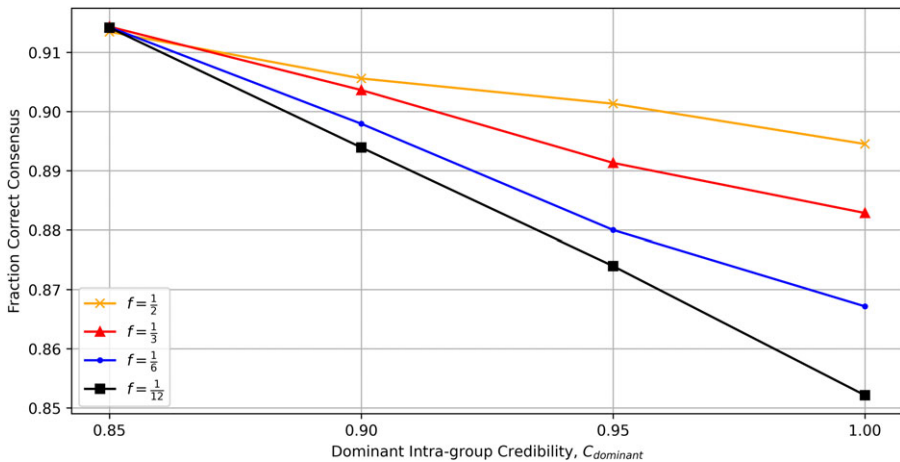
Figure 2. Fraction of simulations reaching correct consensus over a range of $c_{dominant}$ and f values, when $N = 12$, $n = 10$, and $\varepsilon = 0.01$.

dominant group gets larger, with average rounds to correct consensus decreasing as f decreases for all combinations of parameters.

In explaining these results, it is useful to recall Zollman's (2007) observation that, when connectivity in an epistemic community increases, accuracy is sacrificed for speed. A similar tradeoff is at play here. Credence changes that result from updating on reported outcomes are amplified when credibility is over-ascribed. In non-baseline scenarios, members of the dominant group then over-react to reported outcomes provided by dominant group members. When these reports accurately reflect the superiority of action B, dominant group members achieve higher credence values earlier. This increases the speed with which correct consensus is typically achieved. When agents are unlucky and these reported outcomes suggest that action A is superior instead, dominant group members are quicker to lower their credence and abandon the better action. Correct consensus occurs less frequently as a result. Since there are more agents over-reacting to more data as the dominant group gets larger, these effects become more pronounced as f gets smaller.

3.2. Average rounds to threshold credence

As discussed in the previous subsection, when credibility is over-ascribed within the dominant group, correct consensus is reached less frequently, but is typically achieved in fewer simulation rounds. The cost associated with this tradeoff of accuracy for speed is one borne by the entire community – with all agents reaching, or failing to reach, high credence in the relevant true proposition, regardless of group membership. In this subsection, we will examine the distribution of the related speed benefit among members of the two groups.

Simulations are taken to end in correct consensus when all agents have credence values concerning the superiority of action B above 0.99. Measuring how quickly agents in the dominant group and marginalized group reach this threshold provides a sensible way of comparing their relative speed in achieving epistemic success. In order to avoid noisy comparisons involving values very close to the lower limit of 1, only N, f, ε, and n values where the baseline average rounds to correct consensus was above 1.1 will be considered in this subsection.
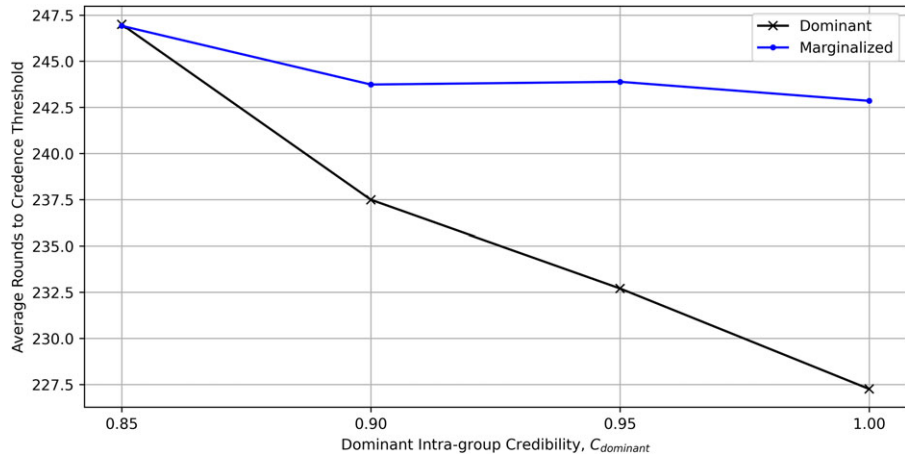
Figure 3. Average rounds to 0.99 credence for agents in the dominant and marginalized groups over a range of $c_{dominant}$ values, when $N = 24$, $f = \frac{1}{2}$, $n = 1$, and $\varepsilon = 0.02$.

Figure 3 shows typical simulation results for the average rounds to threshold credence as $c_{dominant}$ is varied from the baseline value of 0.85 to the maximally over-ascribed value of 1. The major takeaway, seen for all parameter values considered, is that agents in the dominant group reach threshold credence faster than agents in the marginalized group, with this difference increasing as $c_{dominant}$ increases. While average rounds to threshold credence trends down in all cases for the dominant group as the over-ascription of credibility increases, there is no clear trend across parameter values in the speed with which agents in the marginalized group reach this same threshold.

A straightforward explanation can be given as to why this occurs. Recall that, in non-baseline simulations, members of the dominant group update more aggressively on reported outcomes provided by dominant group members. Members of the dominant group will then tend to front-run members of the marginalized group when it comes to adjusting their credence based on general trends in evidence – these trends typically reflect the superiority of action B in this model.[13] When correct consensus is eventually reached, members of the dominant group will then tend to reach high credence values sooner. This is despite the fact that the exact same information is being reported to all members of the community.

Consistent with this explanation, the relative speed advantage enjoyed by dominant group members gets larger as the relative size of the dominant group gets larger (i.e. as f gets smaller). When there is a higher fraction of agents in the dominant group, there is typically more information available per round to aggressively update on, increasing the observed effect.

3.3. Average credence

While average rounds to threshold credence measure how quickly agents in both the dominant and marginalized groups come to have high confidence that the proposition in question is true, it is also worth examining how credence values differ over the entire learning period. To this end, the time-weighted average credence for simulations ending

---

[13]One notable exception occurs when only marginalized group members are performing action B, in which case all relevant attributions of credibility are correct.
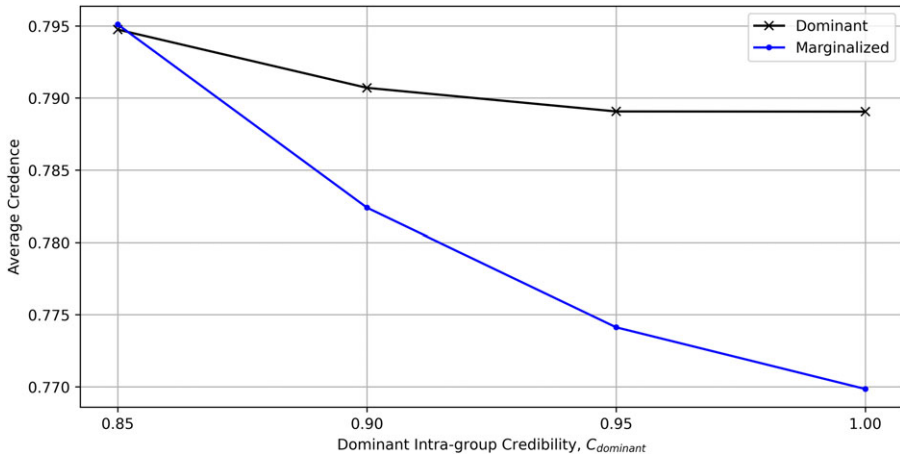
Figure 4. Average credence for agents in the dominant and marginalized groups over a range of $c_{dominant}$ values, when $N = 48$, $f = \frac{1}{12}$, $n = 1$, and $\varepsilon = 0.05$.

in correct consensus was calculated for members of both groups.[14] As in the previous subsection, only N, f, ε, and n values where the baseline average rounds to correct consensus was above 1.1 will be considered.

Figure 4 shows typical average credence values for agents in both the dominant group and marginalized group as $c_{dominant}$ is varied from the baseline value of 0.85 to the maximally over-ascribed value of 1. As can be seen, dominant group agents typically have higher average credence values than marginalized group agents, with this difference increasing as $c_{dominant}$ increases. While this difference is not large in terms of magnitude, it is consistent for all parameter values considered.

The explanation for this effect is once again that members of the dominant group update more aggressively on reported results provided by dominant group members than updating agents in the marginalized group. Since these reported action results typically push credence values in the correct epistemic direction, members of the dominant group will tend to have higher average credence values when the entire successful learning period is considered. This provides a second way in which members of the dominant group are relatively advantaged by the over-ascription of credibility.

4. Conclusion

When introducing a new concept, it is often prudent to begin with examples where what is being picked out is present in an immediate and obvious way. Fricker's reliance on individual instances of testimonial exchange, like Greenleaf's interaction with Sherwood, in bringing to light what she takes to be a distinctly epistemic form of injustice is then unsurprising. By focusing exclusively on these types of examples, however, more concealed forms of the same phenomena – ones that may be no less prevalent or pernicious – can easily be overlooked.[15] This is where the formal tools provided by network epistemology have a unique role to play. By modeling the impact of the prejudicial misallocation of credibility on large groups of individuals engaged in multiple

---

[14]Initial credence values are excluded from this average to avoid unnecessary noise in the calculation.

[15]Certain forms of institutional racism have this same hidden quality. See, for instance, Hardimon's (2020) discussion of "pure" forms of institutional racism where personal racism is absent.

testimonial exchanges that take place over an extended period of time, systematic biases can be uncovered that are largely opaque at the level of individual interactions.[16]

What the simulation results presented in this paper show is that when credibility is over-ascribed within a dominant group, members of that group enjoy systematic epistemic advantages over non-group members. By updating more aggressively on reported outcomes, dominant group members tend to reach high credence values earlier in successful learning periods and enjoy higher credence values on average during those periods, than members of the marginalized group. While this benefit is unevenly distributed, the associated epistemic cost is one borne equally by all community members – with the frequency with which the entire community reaches correct consensus decreasing as the over-ascription of credibility within the dominant group increases. Since epistemic subjects in the marginalized group are then disadvantaged relative to members of the dominant group, the prejudicial over-ascription of credibility within the dominant group should be taken to involve an epistemic injustice.

There are some limitations on the analysis presented in this paper that may warrant further investigation. There could potentially be a level of credibility over-ascription beyond which agents in the dominant group no longer enjoy the epistemic benefits discussed.[17] By running simulations involving r values below those considered in this paper, more extreme levels of over-ascription could be examined.[18] Running simulations involving f values above $\frac{1}{2}$ may also reveal that the epistemic advantages enjoyed by the dominant group are severely curtailed, or even eliminated, when the relative size of the dominant group becomes small.[19] Investigating these possibilities is a task that will be left for future research.

There are two broader methodological points worth making regarding the relevance of the model examined in this paper. First, it could potentially be objected that, rather than being subject to an epistemic injustice, the marginalized group is itself responsible for the epistemic advantage enjoyed by the dominant group. If members of the marginalized group were to update their credence not based on their actual beliefs regarding the reliability of the agents involved, but rather using a method that mirrors how dominant group members with different beliefs update, the relative advantage of the dominant group would disappear, and the two groups would be on equal epistemic footing.

The problem with this suggestion is that it trades one epistemic injustice for another. By updating in a way that is not commensurate with their beliefs, marginalized group members would be behaving irrationally in the standard Bayesian sense – their vulnerability to a diachronic Dutch book reflecting this fact. Forcing the marginalized group to engage in this type of irrational updating in order to counteract the epistemic advantage gained by the dominant group seems itself to be an epistemic injustice.

---

[16]Wu (2023) provides a recent example of how the tools of network epistemology can be used to investigate the hidden effects of prejudice within an epistemic community. What Wu's work uncovers is the epistemic benefits that can be gained by a marginalized group when their testimony is ignored or treated as uncertain by the dominant group.

[17]The greater the level of credibility over-ascription within the dominant group, the greater the chances that the dominant group will be misled by unlucky outcomes that occur within the dominant group. This suggests that there may be such a limit.

[18]Lowering r is necessary because 1 provides the rational upper limit on assigned credibility. The choice of r also impacts the difficulty of the decision problem faced by agents, but so too does the choice of ε and n.

[19]Decreasing the relative size of the dominant group increases the chances that the collection of reported outcomes provided by the dominant group will be misleading.

Second, epistemic models of the type used in this paper almost always involve simplifying assumptions.[20] For instance, it is unlikely that there are real-world groups where every member assigns credibility identically, or real-world epistemic communities where priors are uniformly randomly distributed. By examining artificial scenarios that are simpler than those encountered in real life, features of epistemic communities can be revealed in particularly perspicuous ways.

In this case, a mechanism was identified by which agents who over-ascribe credibility epistemically "front-run" other agents who are more accurate in their credibility ascription. While the epistemic advantages that result will disappear if all agents ascribe credibility identically, this same mechanism should still be operative to some degree as long as the marginalized group is less biased in their ascription of credibility than the dominant group. This is because the credence updates of members of the dominant group will still be larger in magnitude than the credence updates of members of the marginalized group, despite the same information being shared. In addition to providing a "how-possibly" account that is sufficient for defending the thesis that credibility excess can result in cases of epistemic injustice, the model presented is then also likely to provide some insight into how real-world epistemic communities function, even if these communities are more complex and internally varied than those examined in simulation.[21]

In closing, it should be noted that while the modeling framework introduced in this paper was used to examine cases involving the prejudicial over-ascription of credibility within a dominant group, its potential application is much more general. This same framework, which allows for both the over- and under-ascription of credibility within and across groups, could be used to model a wide range of scenarios involving credibility mis-ascription. For instance, the epistemic effects of "tokenism," where members of a marginalized group are given excess credibility concerning an issue related to their social identity, could be modeled.[22] So too could cases of "internalized racism," where members of a marginalized group under-ascribe credibility to themselves, as well as Fricker's (2007) original case of epistemic silencing, where members of a dominant group under-ascribe credibility to members of a marginalized group.[23] Finally, the impact of more widely discussed psychological biases, such as in-group bias and overconfidence bias, on the epistemic success of members of a community could be examined. Future research will focus on using the tools developed in this paper to analyze these and other potential cases of credibility mis-ascription.[24]

References

Bala V. and Goyal S. (1998). 'Learning from Neighbours.' *The Review of Economic Studies* **65**(3), 595–621.

Coady D. (2010). 'Two Concepts of Epistemic Injustice.' *Episteme* **7**(2), 101–113.

Coady D. (2017). 'Epistemic Injustice as Distributive Injustice.' In *The Routledge Handbook of Epistemic Injustice*. Routledge.

David E.J.R., Schroeder T.M. and Fernandez J. (2019). 'Internalized Racism: A Systematic Review of the Psychological Literature on Racism's Most Insidious Consequence.' *Journal of Social Issues* **75**(4), 1057–1086.

---

[20]This is a feature these models share with many standard scientific models. See Weisberg (2013) and Downes (2020) for philosophically interesting discussions.

[21]For other 'how-possibly' defenses of epistemic models, see Pöyhönen (2017), O'Connor and Weatherall (2020), and Wu (2023).

[22]See Davis (2016) and Lackey (2018) for more detailed discussions of this potential form of epistemic injustice.

[23]See David et al. (2019) for a recent survey of the psychological literature on internalized racism.

[24]*Acknowledgements:* I would like to thank Thomas Barrett, Rick Lamb, Alex LeBrun, and an anonymous referee for their valuable comments and suggestions.

Davis E. (2016). 'Typecasts, Tokens, and Spokespersons: A Case for Credibility Excess as Testimonial Injustice.' *Hypatia* **31**(3), 485–501.

Downes S. (2020). *Models and Modelling in the Sciences: A Philosophical Introduction*, 1st edn, New York, NY: Routledge.

Frey D. and Šešelja D. (2020). 'Robustness and Idealizations in Agent-Based Models of Scientific Interaction.' *The British Journal for the Philosophy of Science* **71**(4), 1411–1437.

Fricker M. (1998). 'Rational Authority and Social Power: Towards a Truly Social Epistemology.' *Proceedings of the Aristotelian Society* **98**, 159–177.

Fricker M. (2007). *Epistemic Injustice: Power and the Ethics of Knowing*. New York: Oxford University Press, Oxford.

Fricker M. (2017). 'Evolving Concepts of Epistemic Injustice.' In *The Routledge Handbook of Epistemic Injustice*. Routledge.

Grim P., Singer D.J., Fisher S., Bramson A., Berger W.J., Reade C, Flocken C. and Sales A. (2013). 'Scientific Networks on Data Landscapes: Question Difficulty, Epistemic Success, and Convergence.' *Episteme* **10**(4), 441–464.

Hardimon M.O. (2020). 'Institutional Racism and Individual Responsibility.' In *The Routledge Handbook of Collective Responsibility*. Routledge.

Holman B. and Bruner J. (2017). 'Experimentation by Industrial Selection.' *Philosophy of Science* **84**(5), 1008–1019.

Irzik, G. and Kurtulmus F. (2021). 'Distributive Epistemic Justice in Science.' *The British Journal for the Philosophy of Science*.

Kummerfeld E. and Zollman K.J.S. (2016). 'Conservatism and the Scientific State of Nature.' *British Journal for the Philosophy of Science* **67**(4), 1057–1076.

Lackey J. (2018). 'Credibility and the Distribution of Epistemic Goods'. In M. Kevin (ed.), *Believing in Accordance with the Evidence: New Essays on Evidentialism*, Springer Verlag.

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.

Medina J. (2013). *The Epistemology of Resistance: Gender and Racial Oppression, Epistemic Injustice, and Resistant Imaginations*. New York, Oxford: Studies in Feminist Philosophy, Oxford University Press.

Minghella A. and Highsmith P. (2000). *The Talented Mr. Ripley: A Screenplay*. Methuen, London.

Morehouse K.N. and Banaji M.R. (2024). 'The Science of Implicit Race Bias: Evidence from the Implicit Association Test' *Daedalus* **153**(1), 21–50.

O'Connor C. and Weatherall J.O. (2018). 'Scientific polarization.' *European Journal for Philosophy of Science* **8**(3), 855–875.

O'Connor C. and Weatherall J.O. (2019). *The Misinformation Age: How False Beliefs Spread*. New Haven, CT, USA: Yale University Press.

O'Connor C. and Weatherall J.O. (2020). 'Conformity in Scientific Networks.' *Synthese* **198**(8), 7257–7278.

Pöyhönen S. (2017). 'Value of cognitive diversity in science.' *Synthese* **194**(11), 4519–4540.

Rosenstock S., O'Connor C. and Bruner J. (2017). 'In Epistemic Networks, is Less Really More?.' *Philosophy of Science* **84**(2), 234–252.

Weatherall J.O. and O'Connor C. (2021). 'Endogenous epistemic factionalization.' *Synthese* **198**(25), 6179–6200.

Weisberg M. (2013). *Simulation and Similarity: Using Models to Understand the World*. New York: Oxford Studies in Philosophy of Science, Oxford University Press.

Wu J. (2023). 'Epistemic Advantage on the Margin: A Network Standpoint Epistemology'. *Philosophy and phenomenological research* **106**(3), 755–777.

Zollman K.J.S. (2007). 'The Communication Structure of Epistemic Communities.' *Philosophy of Science* **74**(5), 574–587.

Zollman K.J.S. (2010). 'The Epistemic Benefit of Transient Diversity.' *Erkenntnis* **72**(1), 17–35.

Keith Dyck is a PhD candidate in philosophy at the University of California, Santa Barbara. His research interests include epistemology and philosophy of science.