Epistemic Advantage on the Margin

A Network Standpoint Epistemology

Jingyi Wu*

April 26, 2022
Forthcoming in *Philosophy and Phenomenological Research*.
Please cite the published version.

Abstract

I use network models to simulate social learning situations in which the dominant group ignores or devalues testimony from the marginalized group. I find that the marginalized group ends up with several epistemic advantages due to testimonial ignoration and devaluation. The results provide one possible explanation for a key claim of standpoint epistemology, the inversion thesis, by casting it as a consequence of another key claim of the theory, the unidirectional failure of testimonial reciprocity. Moreover, the results complicate the understanding and application of previously discovered network epistemology effects, notably the Zollman effect (Zollman, 2007, 2010).

> What people of color quickly come to see—in a sense, the primary epistemic principle of the racialized social epistemology of which they are the object—is that they are not seen at all.
>
> Mills (2007, 17)

1 Introduction

Social epistemologists working on race or gender have written extensively on dominant social groups' widespread practice of ignoring or devaluing testimony arising from marginalized groups. For example, Dotson (2011) uses *epistemic quieting* to describe situations in which an audience, often from a dominant social background, fails to identify a speaker, often from a marginalized

*Department of Logic and Philosophy of Science, UC Irvine. Jingyi.Wu@uci.edu.

background, as a knower. Other forms of this practice include *epistemic smothering* (Dotson, 2011), *testimonial injustice* (Fricker, 2007), and a form of *white ignorance* (Mills, 2007). Central to all these cases is a failure of testimonial reciprocity between a speaker and an audience.[1] Moreover, this failure of reciprocity is often unidirectional because members of marginalized groups cannot afford to engage in the devaluation and ignoration of testimony from the dominant group, due to sociopolitical power imbalance (c.f. Mills, 2007, Section 2). I call the situation in which the dominant group ignores testimony from the marginalized group *one-sided testimonial ignoration*, and the situation in which the dominant group devalues testimony from the marginalized group *one-sided testimonial devaluation*. Together, they constitute a unidirectional failure of testimonial reciprocity. That such situations occur is widely claimed in social epistemology, and some authors (e.g. Mills (2007), Saint-Croix (2020), and Wylie (2003)) regard the unidirectional failure of testimonial reciprocity as a key claim of standpoint epistemology, which is a strand of social epistemology that takes as epistemically salient the social positions (or standpoints) that knowers are situated in.[2]

Besides the unidirectional failure of testimonial reciprocity, another key claim of standpoint epistemology, most notably advocated by Hartsock (1983), contends that "certain [socially marginalized] locations *themselves* foster more accurate beliefs, not only concerning one's own social position, but also the social and natural world more broadly" (Saint-Croix, 2020, 493, emphasis in the original). This claim is often called the *inversion thesis*, after the inverse relation between knowers' sociopolitical power and epistemic privilege. The interpretations of and justifications for the inversion thesis are often highly contested (see, e.g., Intemann (2010), Toole (2020), and Wylie (2003)). In what follows, I propose a possible mechanism that gives rise to the inversion thesis, by

[1]For more on testimonial reciprocity, see Hornsby (1995).

[2]There are two other distinct forms of epistemic marginalization discussed in the standpoint epistemology literature that are worth mentioning here. First, one might think that sometimes marginalized groups are not even included in the epistemic community in such a way that they can provide testimony/evidence. Second, one might think that sometimes marginalized groups do not have access to dominant groups' testimony/evidence at all (Narayan, 1988). Interestingly, the base model on one-sided testimonial ignoration that I will present in the paper can be reinterpreted to model this first alternative form of epistemic marginalization as well. This is because the asymmetry in evidence updating dynamics in the model can be interpreted both as marginalized agents providing testimony that is subsequently ignored by the dominant group, *and* as marginalized agents not providing testimony to dominant agents at all (either due to unwillingness or due to epistemic exclusion). These are two very different forms of epistemic marginalization, which interestingly share the same structural form. The models I present in this paper unfortunately do not apply to the second alternative form of epistemic marginalization. I address this limitation in more detail in §5 and leave the additional modeling work for future research. Thanks to an anonymous reviewer for raising these points.

connecting it to the other key claim mentioned above. Specifically, I ask, is it the case that simply by virtue of their testimony being ignored or devalued, members of the marginalized group gain epistemic advantages that foster more accurate beliefs?

I use computer simulations to investigate this question. In my models, members of the marginalized group end up with a number of epistemic advantages, by virtue of their testimony being ignored or devalued by the dominant group. Failure of testimonial reciprocity *can* hence render the inversion thesis true. Though the models I use are highly idealized and abstract, I argue that my simulations provide one *possible* explanation for the inversion thesis, by casting it as a consequence of the unidirectional failure of testimonial reciprocity.[3]

I construct three models to support my argument. My models are adapted from the network epistemology framework developed by Bala and Goyal (1998) and introduced to philosophy of science by Zollman (2007).[4] In all previous implementations of the model in philosophy, the network connections are reciprocal, meaning that if agent Y updates on agent Z's evidence, then Z updates on Y's evidence in the same fashion. In contrast, the network connections in my models are not reciprocal when agents interact with outgroup members.

I start with a base model of one-sided testimonial ignoration. Here, dominant agents ignore testimony shared by marginalized agents, but marginalized agents update on all testimony shared with them.[5] I find that marginalized agents arrive at the true belief more frequently and faster, and select epistemically better actions during the learning process, as compared to dominant agents. Moreover, marginalized agents arrive at the true belief more frequently even compared to agents in a model with perfect testimonial reciprocity, and dominant agents do so less frequently. The entire community in this model learns the true belief less frequently and more slowly than the community with perfect testimonial reciprocity. These results show that the one-sided testimonial ignoration practiced by the dominant group is epistemically detrimental to its members and to the entire community, but is

[3]See Bokulich (2014) for discussions on how-possibly explanations and how-actually explanations. Moreover, though my results may inform real world processes, I do not claim that unidirectional failure of testimonial reciprocity necessarily underlies *all* real world scenarios where the inversion thesis holds. I leave open the possibility that other phenomena may also lead to the inversion thesis. That is, I provide a sufficient condition for the inversion thesis under reasonable assumptions, but not a necessary condition.

[4]Variations of this model have seen fruitful applications in the philosophy of science and social epistemology, e.g. Zollman (2007), Zollman (2010), Mayo-Wilson et al. (2011), and O'Connor and Weatherall (2018).

[5]I use "marginalized agents" to denote "members of the marginalized group," and "dominant agents" to denote "members of the dominant group."

epistemically advantageous to the marginalized group.

I then construct two variations of the base model—one for one-sided testimonial ignoration with the homophilic network structure and one for one-sided testimonial devaluation. I use homophilic networks—where agents prefer to connect with ingroup members—because many real world networks are homophilic (McPherson et al., 2001) and because these networks allow me to vary agents' information access based on group membership. I find that, regardless of their information access, members of the marginalized group arrive at the true belief more frequently than the dominant group in homophilic networks. The degree to which marginalized agents gain other epistemic advantages, such as their speed of learning, depends on their information access. Finally, I build a model of one-sided testimonial devaluation, where dominant agents discount, rather than ignore, testimony from the marginalized group. Here, marginalized agents arrive at the true belief faster and select epistemically better actions during the learning process, as compared to dominant agents.[6]

The paper will be organized as follows. §2 introduces and motivates the base model, as well as presents the key results. §3 discusses the first variation with homophilic network structures. §4 presents the second variation on testimonial devaluation. §5 discusses how my results relate to standpoint epistemology. In closing, I will briefly note how my models, by virtue of introducing nonreciprocity to network connections, complicate the understanding and applications of previous network results.

2 The Base Model: one-sided testimonial ignoration

2.1 The Model

The base model consists of a network of agents who are presented with the same learning problem. Agents are tasked with learning which of the two available options is better, by updating on evidence from their neighbors and themselves. The network has two subgroups—the marginalized group and the dominant group—with their members distinguished only by the updating dynamics. Dominant agents only update on evidence shared by ingroup members, whereas marginalized agents update on

[6] Due to model design, the marginalized and dominant groups necessarily learn the true belief with the same frequency.

evidence shared by everyone.

2.1.1 The Bandit Problem

To motivate my model, let us consider a toy example. The scenario is not meant to be realistic, but rather to illustrate the cases to which my models are applicable in a high-level way. Suppose that an organization hires for a position, and eventually offers the position to a candidate from a particular social group X.[7] Suppose further that this is not the first time that a candidate from X is hired, and Hana, a consultant who has access to some details of the case, is tasked to investigate why the candidate was hired. There is an available individualistic meritocratic explanation, A, according to which a candidate from X is hired because they are the best at doing the job out of all candidates. Explanation A is well understood, but is only known to succeed about half the time when applied to similar cases. For instance, there might have been multiple candidates who were equally good at the job but only one was hired. Explanation A is all right, but is inadequate as a catch-all explanation. Recently, a new explanation called structural bias explanation, B, is also hypothesized to account for this kind of social phenomena. Explanation B says that a candidate from social group X is hired because during the hiring process there was structural bias against candidates who are not members of X. Explanation B is not well understood, and the community is unsure whether it is better or worse than A. Suppose that Hana happens to have the initial belief that B would be better than A in this case, so she decides to solicit evidence to test B by, for instance, investigating whether the company's job description contains biased language.[8] She then learns from the evidence she discovers, forms posterior beliefs on the explanations, and continues to test B if she thinks that it is better. If she keeps getting good evidence for B, eventually she should believe that B is better than A with overwhelming confidence. Or, if the evidence Hana gets for B is unsatisfactory, she would give up on testing B because she thinks it is inferior to the available alternative.

The above scenario can be modeled by what is called a two-armed bandit problem. The name "bandit problem" comes from applying it to a gambling scenario, where a gambler, facing a many-armed "bandit," aims at maximizing their profit and choosing the best-performing arm by interacting

[7]To make it feel more concrete , readers can substitute "X" with "White," "Male," "Able-bodied," etc.

[8]Here we stipulate that, depending on what hypotheses to test, Hana performs different actions, which then provide Hana with evidence for the chosen hypothesis. Furthermore, Hana has limited resources to test explanations, so she is incentivized to test the better explanation each round. See §5 for potential limitations of this stipulation.

with the machine. Here is how the problem is set up for one agent. For every time step (or "round"), the agent selects between two actions: A and B, and gets payoffs based on their choice. Each of the actions is associated with a fixed probability of success. The success rate for A is well-known to the agent, and is set to $.5$. The success rate for B, however, is uncertain to the agent: the agent knows that action B is either slightly better than A, with a success rate of $.5 + \epsilon$, or it is slightly worse than A, at $.5 - \epsilon$. When an action generates success, the agent receives a payoff of 1, and they receive no payoffs otherwise. In the models I implement, unbeknownst to the agents, I set the success rate of B to $.5 + \epsilon$. The goal for the agent is to determine which action has a higher success rate by learning from their own actions and payoffs.[9]

Thus constructed, the two-armed bandit problem is suitable to model learning situations where there are two competing choices. This model is often applied in epistemology to scenarios with two competing theories, explanations, or hypotheses available for a given phenomenon. Besides Hana's quest to figure out which of the two explanations best accounts for the company's hiring decision, another classic application of the model is the clinical trials of drugs. Here, action A represents a drug that is well-understood, and the doctor's goal is to figure out whether a new drug, B, is better or worse than A.

I use the bandit problem because we can set the true state of the world one way and observe how well the agent learns the true belief, which naturally models epistemic advantage. Moreover, when we cast the bandit problem in a social setting, the evidence sharing dynamic becomes a suitable place to implement the unidirectional failure of testimonial reciprocity, as I will discuss shortly.

2.1.2 Going Social

In many real cases, learning is not an isolated, individualistic activity. Doctors are often not alone when they test new drugs; they may be in contact with other doctors in the same clinical trials. Hana may also have a team of consultants on the same case.

When multiple individuals figure out the same problem together, they can share evidence and incorporate others' evidence in their own learning. To model this, I introduce a network of agents who face the same two-armed bandit problem. Each agent is connected to some or all of the other

[9] Agents learn by applying Bayes' rule. For more detail, see §2.1.4.

agents and I call the agents they are connected to their "neighbors."[10] In each round, each agent selects their action based on their belief in the proposition "B is better than A," obtains evidence from their action, shares their evidence with their neighbors, and updates on the evidence they receive. Further, because the evidence is passed on between agents, a piece of evidence is a form of testimony—it has a speaker and an audience.

Zollman (2007) builds a model just like this. He simulates the model on different network structures and finds, perhaps counterintuitively, that a more sparsely connected community has epistemic advantages over a more connected one, in the sense that the former learns the true belief more frequently. However, a more sparsely connected community also learns the truth more slowly. Zollman's findings are collectively dubbed "the Zollman effect." They uncover a trade-off between the speed and accuracy in social learning. The effect will become relevant later in two ways. First, I will use the reasoning behind the effect to explain some of my modeling results. Second, I will argue that my results complicate the interpretation and application of the Zollman effect.

2.1.3 De-idealizing Testimonial Relationships

In Zollman (2007)'s model, if two agents are connected, then they share their evidence with each other and fully update on the evidence they receive. That every agent treats all testimony they receive equally is, of course, an idealization, and perhaps an unwarranted one. Social epistemologists working on race and gender, notably standpoint epistemologists, have written extensively on dominant social groups' widespread practice of silencing testimony from marginalized perspectives.

Recall the epigraph of this paper. Mills (2007) argues that the primary epistemic principle of a racialized social epistemology is that people of color are not seen as knowers. In the same chapter, Mills gives the example of Kant's dismissal of a Black carpenter's epistemic credibility: "and it might be, that there were something in this which perhaps deserved to be considered; but in short, this fellow was *quite black* from head to foot, a clear proof that what he said was stupid" (Kant, 1960, 113, emphasis in the original, qtd. Mills 2007, 32). Moreover, several authors point out that people of color's testimony is often not taken seriously unless they have white authenticators (Bright, forthcoming; Fatima, 2017; Mills, 2007), and women's testimony is often ignored until repeated by

[10] A *network structure* describes how agents are connected to each other.

men, a phenomenon dubbed "hepeating" (Deo, 2019). It is hypothesized that Ida B. Wells-Barnett (2012), when arguing against lynching in 1895, only includes evidence from white sources because her intended white audience would trust these sources rather than the testimony of Black people (Bright, forthcoming).[11] As a contemporary example, Deo (2019) conducts an empirical study on how race and gender influence legal academia,[12] and finds that "most women in the [study] sample, regardless of racial/ethnic background, have endured silencing, harrassment, mansplaining, hepeating, and gender bias" (Deo, 2019, 47). For instance, one study participant has "counted over ten times on [her] faculty where [she has] said something and [nobody has responded; then] a male faculty has repeated it and another male colleague has said, 'Good idea!'" (Deo, 2019, 45). Writing about her experience as a woman of color in the predominantly white and male field of professional philosophy (an experience that corroborates Deo's findings), Fatima (2017, 151) claims that "if the only way that a woman of color's testimony is given any uptake is if dominant members of academia verify it, then we have already discounted the epistemic credibility of the speaker." The existence and prevalence of silencing testimony from marginalized perspectives is widely recognized.

Importantly, this dismissal of testimony is one-sided. Mills (2007, 17) argues that "often for their very survival, blacks have been forced to become lay anthropologists, studying the strange culture, customs, and mind-set of the 'white tribe' that has such frightening power over them, that in certain time periods can even determine their life or death on a whim." Mills quotes Baldwin's brutally honest line, "I have spent most of my life, after all, watching white people and outwitting them, so that I might survive" (Baldwin, 1993, 217, qtd. Mills 2007, 18). In Deo (2019)'s study, a woman of color participant admits that she always acquiesces to requests from university administration, sometimes even unreasonable ones, because she fears professional repercussions if she declines. In order to survive in a hegemony dominated by other people, marginalized folks *cannot afford to* ignore testimony and demands from the dominant group.

Given this, it is appropriate to de-idealize the testimonial relationships in the model to incorporate one-sided testimonial ignoration.[13] I implement testimonial ignoration by dividing the population

---

[11] Though Bright (forthcoming) eventually favors an alternative, statistically-based explanation for Wells-Barnett's decision, the original hypothesis is still plausible as testimonial ignoration was undoubtedly salient at the time.

[12] In this comprehensive study, Deo (2019) presents quantitative and qualitative results from a core sample comprising almost 10% of all US women of color law professors, together with a comparison sample of white or men of color law professors.

[13] I will not address *how* we identify situations with one-sided testimonial ignoration. Dotson (2011) has gracefully

into two groups: the marginalized group and the dominant group. Marginalized agents update on evidence from all their neighbors, but dominant agents only update on evidence shared by ingroup neighbors.[14] Here, we have a failure of testimonial reciprocity—marginalized agents take testimony from the dominant group as they are meant to be taken, but dominant agents fail to do so reciprocally. To bring us back to the hiring scenario, there might be a few consultants who do not trust Hana's evidence, as (supposedly) Hana is not a member of the social group X and, for them, Hana might have bias against X.[15]

2.1.4 Further Technical Details

A few technical details are in order before I present my results.

Initiation

I set the success rate for A to be $.5$, and the success rate for B to be $.5 + \epsilon$. At the start of the simulation, every agent is assigned a credence randomly selected from a uniform distribution between 0 and 1 (exclusive). The credence reflects their belief in the proposition "B is better than A."

A Typical Round

At the start of every round, each agent selects one of the two actions—if their credence is $> .5$, they choose action B; otherwise they choose A. The agent then performs the chosen action a number of times, n, and receives payoffs.

tackled this question.

[14] One might worry that the evidence I presented in the previous paragraph only supports the claim that *insofar as* marginalized agents receive evidence from dominant agents, they cannot afford to ignore the evidence, but not so much that marginalized agents receive all the evidence from their dominant neighbors. This is a very fair concern, especially given that sometimes marginalized knowers are excluded from participating in epistemic communities (more in §5). The base model I present in this section makes the idealizing assumption that marginalized agents receive evidence from all their neighbors. However, one can reinterpret the first variation presented in §3 with homophilic networks, where p_{outgroup} is small, as modeling some version of epistemic exclusion. Here, marginalized agents only have very sparse evidential access to the dominant group. We can think of this as the poignant situation where a small number of marginalized knowers are invited to participate in dominant epistemic spaces, but their testimony is still ignored. Thanks to an anonymous reviewer for raising this concern.

[15] For a more realistic example, consider Blanche, a Black fill-in maid for a white family in the novel *Blanche on the Lam* (Neely, 1993). As a lower class Black woman, Blanche's epistemic credibility is not fully recognized by other members of the family that employs her; but Blanche continues to listen to and in on the family members during her work. This eventually allows Blanche to gather enough evidence and solve a series of murder mysteries in the family. This example of Blanche's standpoint is discussed at length in Wylie (2003). Though the example of Blanche is fictional, the phenomenon that Black domestic helpers have their epistemic credibility suppressed by their employers, but still gain an outsider-within status in white middle-class families is discussed in detail in Collins (2002).

Then, each agent uses Bayes' rule to update their credence based on both their own experience and the experiences of their neighbors. For example, suppose that $\epsilon = .1$ (i.e. the success rate of B is .6). If an agent has prior credence of .7 that B is better than A, and pulls action B one time, which succeeds in generating a payoff of 1, then their posterior credence after updating on their own experience is

$$P(H|E) = \frac{P(E|H)P(H)}{P(E|H)P(H)+P(E|\neg H)P(\neg H)} = \frac{.6\times.7}{.6\times.7+.4\times.3} = .78.$$

Here, H (hypothesis) stands for "B is better than A," and E (evidence) is "taking action B once yields 1 payoff." It is worth noting that, whether or not the action succeeds in generating a payoff, performing action A will not change the posterior credence. To see that, we observe that $P(E'|H) = P(E'|\neg H) = P(E''|H) = P(E''|\neg H) = .5$, where E' is "taking action A once yields 1 payoff," and E'' is "taking action A once yields 0 payoffs." Consequently, $P(H|E') = P(H|E'') = P(H)$. Agents similarly update on neighbors' evidence by applying Bayes' rule. Note that dominant agents only update on evidence from ingroup neighbors and marginalized agents update on evidence from all neighbors.

After each agent finishes updating, we increase the time step by 1 and repeat the procedure for a typical round.

End of Learning

There are three stable end states for this model:

- Community convergence to the true belief: every agent has a credence of $> .99$ that B is better than A. In this state, it is increasingly unlikely that agents would switch from B to A. Everyone succeeds in learning.
- Community convergence to the false belief: every agent has a credence of $\leq .5$ that B is better than A. In this state, nobody would be actively testing B. Everyone fails in learning.
- Polarization: every marginalized agent has a credence of $> .99$ that B is better than A, and every dominant agent has a credence of $\leq .5$. In this state, no dominant agent would be actively testing B. Every marginalized agent succeeds in learning, every dominant agent fails in learning, and the entire community fails in learning.

Due to one-sided testimonial ignoration, polarization is a new end state for my model compared to Zollman (2007). In this state, even though marginalized agents are still testing action B, their testimony is ignored by dominant agents. We have a stable situation where the agents' beliefs are split along group membership. It is worth noting that a polarization with the opposite distribution of credence cannot be stable, since in this state, marginalized agents would still update on evidence from dominant agents, and the model would evolve. If the network reaches one of the stable end states, the community has finished learning.

2.2 Results

I simulate this model using the following values for the key parameters, for $10,000$ runs each:

- Total number of agents ("size") of the network (k) : $3, 6, 12, 18$.
- Proportion of the marginalized group in the population(d): $\frac{1}{6}$, $\frac{1}{3}$, $\frac{1}{2}$, $\frac{2}{3}$.
- Number of pulls per round (n): $1, 5, 10, 20$.
- Probability of B (P_B): $.51$, $.55$, $.6$, $.7$, $.8$.
- Network structure: complete.[16]

For all parameter values, I run a comparison model with perfect testimonial reciprocity; this is equivalent to the complete model from Zollman (2007). Moreover, all the results in this paper are robust across all parameter values, unless otherwise noted.

I employ three ways to measure how well subgroups learn: the (1) frequency and (2) speed at which subgroups learn the true belief, and the (3) frequency at which subgroup members select the epistemically better action during the learning process. The marginalized group holds epistemic advantages compared to the dominant group according to all three measures.

I measure how frequently a subgroup learns the true belief by calculating the proportion of simulation runs (out of $10,000$ runs) where the subgroup succeeds in learning. This measure captures how often an average agent of a given subgroup eventually learns the true belief. Here, the marginalized group learns the true belief more frequently (Figure 1) because polarization counts as success for marginalized agents and failure for dominant agents. As long as there are simulation runs that end in polarization, the marginalized group would learn better in this respect. In fact, for all parameter values, the marginalized group learns the true belief more frequently.

A subgroup's speed of learning the true belief is measured as follows. For each agent and each of their successful runs, I document the earliest round after which the agent maintains a credence of $>$ $.99$. Call this an individual agent's rounds to successful learning. Then, for each subgroup, I compute

[16] A network is complete when everyone is connected to everyone. The network structure here is complete prior to adding one-sided testimonial ignoration.

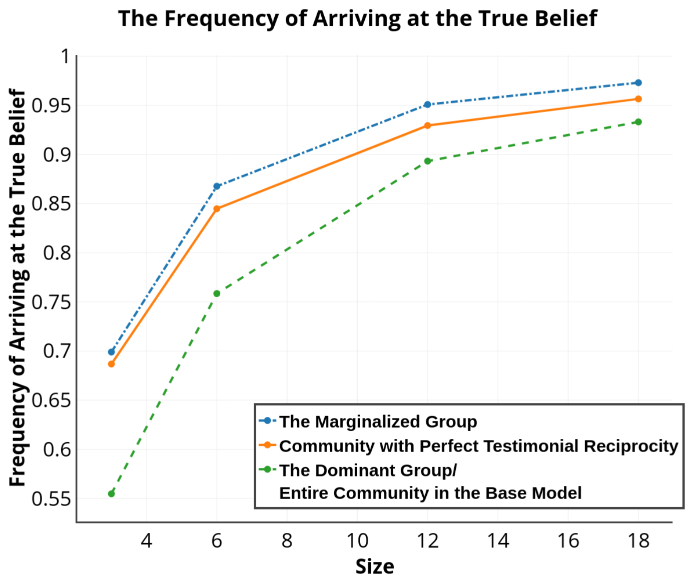

Figure 1: Base Model, $P_B = .6, n = 1, d = \frac{1}{3}$, $10,000$ simulation runs.

the subgroup's average rounds to successful learning by taking the average of all individual agents' rounds to successful learning, out of all successful simulation runs and all agents in the subgroup. Unlike some previous measures of speed of learning, here I only consider cases where the subgroup learns the truth.[17] This measure represents how long an average agent from a given subgroup takes to learn the true belief. Marginalized agents learn the true belief faster (Figure 2) because they have access to more information per round.

I measure how frequently subgroup members select the epistemically better action during the learning process by calculating the average proportion of rounds that an agent of this subgroup selects action B, out of all simulation runs and all agents in the subgroup, without considering every agent's selection at round 0.[18] This measure captures how frequently an average agent from a given subgroup chooses the epistemically better action during the learning process. It is not surprising that marginalized agents select the better action more frequently, since they also eventually learn the truth more frequently and faster.

[17] My measure differs from previous ones in the literature. Different from Zollman (2007)'s "average time to success," I measure a subgroup's speed of learning not by observing when the entire community reaches the true belief, but by taking the average of individual agent's rounds to successful learning. "Rounds to consensus" in O'Connor and Weatherall (2018) measures the entire community's rounds to consensus, regardless of truth or falsity. In contrast, I only consider cases where the subgroup learns the truth. This measure allows me to capture possible differences in the speed of successful learning between the two subgroups. For instance, there could be possible variations in subgroups' rounds to successful learning even when the community converges to the true belief. Later, I will introduce another measure: the *entire community*'s average rounds to successful learning, which is the same as Zollman (2007)'s.

[18] An agent's action at round 0 only depends on their initial credence as randomly selected by a uniform distribution. Including this round would add noise to the data, especially when agents learn very fast (when P_B and n are large).

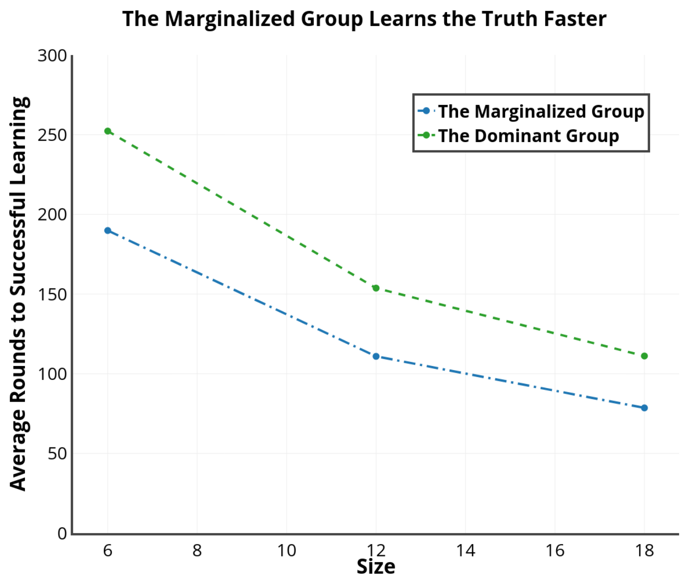

Figure 2: Base Model, $P_B = .51, n = 10, d = \frac{1}{2}$, $10,000$ simulation runs.

So far, my results adhere to a basic empiricist intuition, that access to more information provides epistemic advantages. However, the following results, together with results from the first variation (§3), suggest that marginalized agents' epistemic advantages cannot solely be explained by having access to more information.

Compared with the community with perfect testimonial reciprocity, where all evidence is fully updated by the receiver, the marginalized group in my model learns the true belief more frequently, and the entire community as well as the dominant group learns the true belief less frequently (Figure 1). A marginalized agent holds this epistemic advantage even though they update on as many pieces of evidence as an agent in the model with perfect testimonial reciprocity. This is because in simulation runs that start with initially unpromising results for action B,[19] the community with perfect testimonial reciprocity would quickly settle on action A, resulting in a convergence to the false belief. In my model, though marginalized agents would change their actions quickly, the dominant agents would not due to their limited information access. Consequently, the epistemically better action remains active in the network for longer, making it more likely that marginalized agents would turn around to correct their action. This is an instance of the Zollman effect, where more sparsely connected network structures produce epistemic benefits.[20]

[19]Because whether each action succeeds is probabilistic, these scenarios occur in my simulations.

[20]This result is robust with parameters such that the average rounds to successful learning for the entire community is > 3 (i.e. excluding "easy" learning situations with large P_B and n). One reason for the non-robustness in the edge cases is related to the trade-off between the learning speed and learning accuracy. For cases where the learning is really

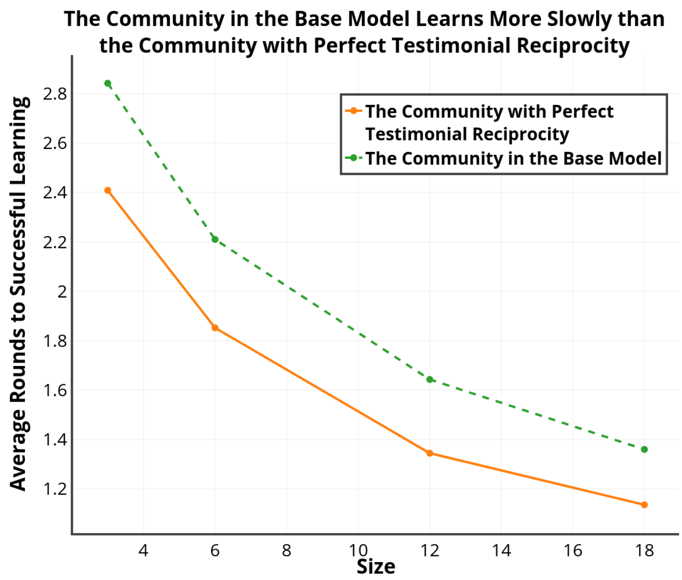

Figure 3: Base Model, $P_B = .7, n = 5, d = \frac{1}{3}, 10,000$ simulation runs.

The entire community in my model learns the true belief less frequently than the community with perfect testimonial reciprocity because a necessary condition for the former community to learn the true belief is that the dominant group learns it. But the dominant group, because of the one-sided testimonial ignoration, acts as an isolated model of size $k \cdot (1 - d)$ with perfect testimonial reciprocity.[21] For models with perfect testimonial reciprocity, the smaller the size of the network, the less frequently the community learns the true belief (Zollman, 2007). The dominant group, and hence the entire community, learns the true belief less frequently than the community of size k with perfect testimonial reciprocity.

Moreover, the entire community in my model learns the true belief more slowly than the community with perfect testimonial reciprocity (Figure 3). The speed of successful learning for the *entire community* is measured differently from the subgroups, to facilitate a direct comparison with Zollman (2007)'s model. I first record, for each simulation run with community success, the round at which the community finishes learning. I then define the average rounds to successful learning for the entire community as the average of these rounds out of all successful simulation runs. This measure captures how long an entire community takes to learn the true belief. The entire community in my model learns the true belief more slowly for a similar reason—the dominant

fast, the speed of successful learning is often very close between the marginalized group and the community with perfect testimonial reciprocity, so their learning accuracy is also comparable. Moreover, when agents finish their learning quickly, their learning accuracy is highly dependent on their initial beliefs, which are randomly assigned and highly variable.

[21] Recall that k is the network size and d is the proportion of the marginalized group.

group, as an isolated group with perfect testimonial reciprocity of size $k \cdot (1 - d)$, learns the truth more slowly than a community with perfect testimonial reciprocity of size k (Zollman, 2007).

Finally, the proportion of the marginalized group (d) impacts the degree of epistemic (dis-)advantage. As d increases, the marginalized group learns the true belief more often,[22] and the dominant group less often. This is because the size of the dominant group decreases as d increases. In the face of unpromising initial results for B, dominant agents are even more less likely to quickly give up on B due to their further limited information access, resulting in more gain in learning accuracy for the marginalized group. Furthermore, as d increases, the entire community learns the truth more slowly for similar reasons.

In this model, I do not restrict the proportion of the marginalized group in the population. What defines a subgroup's marginalized status is the one-sided testimonial ignoration, rather than its size. This is a merit since in many real cases, the marginalized group can be in the majority, such as during the Apartheid in South Africa.

3 Variation 1: Homophilic Networks

3.1 The Model

As mentioned, some results from the based model can be explained by marginalized agents having access to more information, but other results cannot solely be explained by information access. Simulating the model with homophilic networks—where agents prefer to connect with ingroup members—allows me to further investigate the extent to which information access influences epistemic advantages. Moreover, homophily is a natural choice, as many real human networks are homophilic based on race, gender, class, etc. (McPherson et al., 2001). My results show that marginalized agents still hold a number of epistemic advantages, even when they have equal or fewer expected connections compared to dominant agents.

This variation differs from the base model in network structure only. I use two-type random graphs to generate the homophilic networks.[23] First, every agent is connected to themselves. Then, I

[22] This is robust with parameters such that the average rounds to success for the entire community is > 3.

[23] See Golub and Jackson (2012) and Rubin and O'Connor (2018) for previous implementations of homophilic networks using this method.

divide the agents into marginalized and dominant groups. Each agent has some probability of connecting with ingroup members, P_{ingroup}, and some other probability of connecting with outgroup members, P_{outgroup}. Finally, I require that, prior to adding one-sided testimonial ignoration, the network structure is undirected. This means that Y is connected to Z if and only if Z is connected to Y. The network is homophilic when $P_{\text{ingroup}} > P_{\text{outgroup}}$.

When a subgroup is in the minority, its members have fewer expected connections than members of the other group, prior to adding one-sided testimonial ignoration. To see this, first observe that the expected number of connections for an agent in this subgroup is

$$P_{\text{ingroup}} \cdot (k \cdot d' - 1) + P_{\text{outgroup}} \cdot k \cdot (1 - d') + 1,$$

where k is the size of the network, and d' is the proportion of this subgroup in the population. For an agent in the other group, their expected number of connections is

$$P_{\text{ingroup}} \cdot (k \cdot (1 - d') - 1) + P_{\text{outgroup}} \cdot k \cdot d' + 1.$$

When $d' < \frac{1}{2}$ and $P_{\text{ingroup}} > P_{\text{outgroup}}$, an agent in this subgroup has fewer expected connections than an agent in the other group.[24]

When one-sided testimonial ignoration is added and when the marginalized group is in the minority, we can specify the values of P_{ingroup}, P_{outgroup}, and d such that a marginalized agent would have fewer, equal, or more expected connections as compared to a dominant agent. For example, a marginalized agent would have the same number of connections as a dominant agent when $P_{\text{ingroup}} = .8$, $P_{\text{outgroup}} = .4$, and $d = \frac{1}{3}$, fewer expected connections when $P_{\text{ingroup}} = .8$, $P_{\text{outgroup}} = .3$, and $d = \frac{1}{3}$, and more expected connections when $P_{\text{ingroup}} = .8$, $P_{\text{outgroup}} = .5$, and $d = \frac{1}{3}$. Simulating with homophilic networks, then, can reveal the extent to which information access influences marginalized agents' epistemic advantages.

3.2 Results

I simulate this variation using the following values for the key parameters, for $10,000$ runs each:[25]

- Size of the network (k) :18.

[24] The result holds probabilistically and is not necessarily true for individual simulation runs.
[25] I randomly generate a homophilic network for every simulation run.

- Number of pulls per round (n): 1, 5, 10, 20.
- Probability of B (P_B): .51, .55, .6, .7, .8.
- Proportion of the marginalized group (d): $\frac{1}{6}$.
  - $P_{\text{ingroup}} = .8, .9, 1.$
  - $P_{\text{outgroup}} = .6, .7, .8.$
- Proportion of the marginalized group (d): $\frac{1}{3}$.
  - $P_{\text{ingroup}} = .7, .8, .9.$
  - $P_{\text{outgroup}} = .3, .35, .4, .45, .5.$

I choose different values for P_{ingroup} and P_{outgroup} based on d because depending on the value of d, the values of P_{ingroup} and P_{outgroup} needed for a marginalized and a dominant agent to have the same number of connections are different.

Furthermore, I only simulate connected* networks in order to reduce noise in the data.[26] Because of the one-sided testimonial ignoration, I define connectedness* as follows. A network is connected* when (1) there exists a path from any marginalized agent to any arbitrary agent in the network, and (2) there exists a path from any dominant agent to any arbitrary dominant agent. Moreover, there exists a path from agent Y to agent Z iff there are agents $A_0, A_1, ..., A_i$ with $i \geq 1$ in the network such that (1) $Y = A_0$ and $Z = A_i$, and (2) A_k updates on evidence shared by A_{k+1}, with $0 \leq k \leq i-1$. I only test the network size of 18 for two practical reasons. First, the total number of simulations is already large due to variations in P_{ingroup} and P_{outgroup}. Moreover, networks with large populations are more likely to be connected* after the two-type random graph generation process.

My results show that the marginalized group learns the true belief more frequently than the dominant group, regardless of their expected numbers of connections. Similar to the base model, this is due to polarization. When polarization occurs, the marginalized group succeeds in learning but the dominant group fails, creating a disparity in learning accuracy. Moreover, the frequency of learning the true belief for the two subgroups does not change drastically as P_{outgroup} changes (Figure 4). This is because P_{outgroup} does not influence the epistemic behavior of the dominant group as an isolated community. As a result, the epistemic benefits the marginalized group gains remain the same.

[26] If the network is not connected*, then there would necessarily be two or more isolated communities without any evidence sharing in between. This network would produce less than typical learning speed and learning accuracy, compared to connected* counterparts with same parameter values.

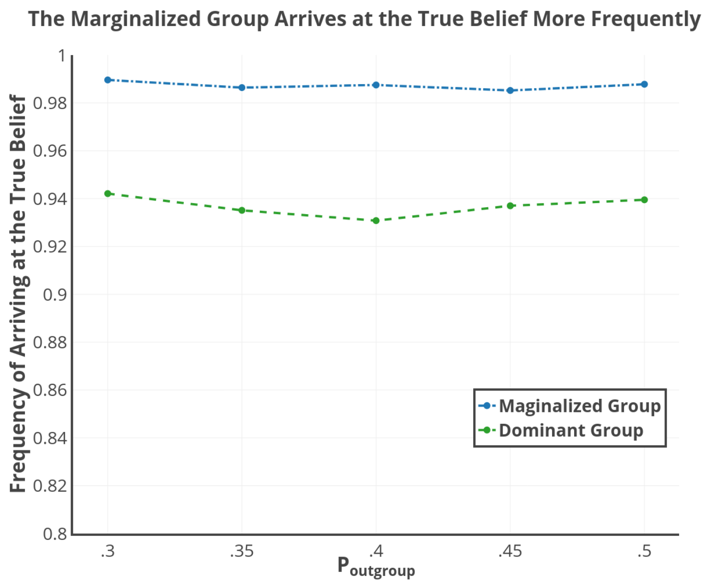

Figure 4: Variation 1, $P_B = .55, n = 5, d = \frac{1}{3}, P_{\text{ingroup}} = .8, k = 18, 10,000$ simulation runs.

How the two subgroups compare regarding their learning speed and the frequency of choosing the epistemically better action depends on their members' number of connections. When a marginalized agent has the same expected number of connections as a dominant agent, the former selects the epistemically better action (i.e. action B) more frequently than the latter during the learning process (Figure 5).[27] However, unlike the base model, marginalized agents in general learns the true belief more slowly than dominant agents.[28]

The reason is that marginalized agents learn much more slowly when simulations end in polarization than when simulations end in community success, but polarization still counts as success for them.[29] If we discount polarization from the marginalized group's average rounds to successful learning, then the marginalized group also learns the truth faster than the dominant group in many but not all cases.

When a marginalized agent has more expected connections than a dominant agent, the marginalized group's epistemic advantage in speed of successful learning becomes more robust, while its members continue to hold the other advantages. As the difference in the expected numbers of connections grows, eventually the average rounds to successful learning for the marginalized

[27]The behaviors of the two subgroups get closer as P_B increases. When P_B is large, the two possible states of the world become easier to distinguish. Therefore, the agents finish quickly, at less than 2 rounds. The frequency of selecting the better action conveys less information as P_B gets large.

[28]Measured in the first way introduced in §2.2.

[29]The marginalized group takes around two to five times more rounds to learn the true belief when simulations end in polarization than those that end in community success.

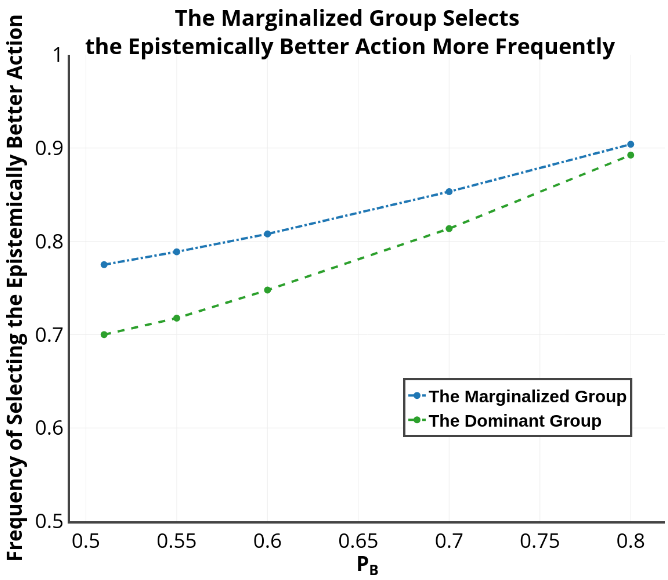

Figure 5: Variation 1, $n = 1, d = \frac{1}{3}, P_{\text{ingroup}} = .9, P_{\text{outgroup}} = .45, k = 18, 10,000$ simulation runs.

group, including polarization, would be lower than that for the dominant group.[30]

When a marginalized agent has fewer expected connections than a dominant agent, the marginalized group loses its epistemic advantage in the speed of successful learning. However, in many cases where the difference in numbers of expected connections is small ($\leq\ 2$), the marginalized group retains its epistemic advantage in the frequency of selecting the epistemically better action during learning. Hence, the epistemic advantages brought to the marginalized group by one-sided testimonial ignoration is sometimes strong enough to offset the potential disadvantages from the loss of information access.

4 Variation 2: one-sided testimonial devaluation

4.1 The Model

I now simulate testimonial devaluation, where dominant agents devalue evidence from the marginalized group, rather than ignore it. The model differs from the base model only in updating rules. Here, I introduce Jeffrey conditionalization, which allows agents to update on evidence according to how much they trust the accuracy of it. The formula for Jeffrey conditionalization is

[30]The base model is a special case of a two-type random graph, where $P_{\text{ingroup}} = P_{\text{outgroup}} = 1$. The results in §2.2 fits with those here.

the following:

$$P_f(H) = P_i(H|E) \cdot P_f(E) + P_i(H|\neg E) \cdot P_f(\neg E).$$

The agent's final credence for the hypothesis H ($P_f(H)$) is defined as the agent's initial credence for H after Bayesian conditioning on the evidence E being true ($P_i(H|E)$) times the agent's credence that E is accurate ($P_f(E)$), plus the the agent's initial credence for H after Bayesian conditioning on E being false ($P_i(H|\neg E)$) times the agent's credence that E is inaccurate ($P_f(\neg E)$). When $P_f(E) = 1$, the agent fully trusts the accuracy of E, and Jeffrey conditionalization reduces to Bayes' rule. When $P_f(E) = P_i(E)$, i.e. the agent's final credence for E equals their initial credence for E,[31] then $P_f(H) = P_i(H)$, i.e. the agent keeps their original credence for H and ignores the evidence altogether. When $P_f(E)$ is between $P_i(E)$ and 1, the agent positively updates on the evidence, though not fully.

O'Connor and Weatherall (2018) use Jeffrey conditionalization to model situations where agents do not fully trust the evidence gathering practices of other agents. In their models, agents' final credence for the evidence $P_f(E)$ is based on how similar the sharer's belief is to the updater's. Agents find the evidence shared by someone with similar beliefs more trustworthy. In my model, however, the distrust is not based on the relative similarity of beliefs, but rather on group membership.

As before, the entire population is divided into the marginalized group and the dominant group. Marginalized agents fully update on evidence shared by all neighbors; dominant agents, in contrast, fully update on evidence shared by ingroup neighbors, but devalue evidence shared by outgroup members by applying Jeffrey conditionalization, with $P_f(E)$ calculated using:

$$P_f(E) = 1 - m \cdot (1 - P_i(E)).$$[32]

Here, m is a parameter between 0 and 1, capturing how much dominant agents devalue testimony from the marginalized group. When $m = 0$, $P_f(E) = 1$ and dominant agents fully update on evidence from marginalized agents—the model is equipped with perfect testimonial reciprocity.

[31] $P_i(E)$ can be calculated from $P_i(H)$ in the following way: $P_i(E) = P_i(E|H)P_i(H) + P_i(E|\neg H)P_i(\neg H)$.

[32] There are several formulae for $P_f(E)$ that would satisfy the desiderata below equally well. However, my results would remain largely the same had I chosen the alternatives. Furthermore, because Jeffrey conditionalization is non-commutative (c.f. Lange, 2000), I require that a dominant agent randomly selects the order according to which they update. The order of updating, to my knowledge, does not influence the qualitative results.

When $m = 1$, $P_f(E) = P_i(E)$ and dominant agents fully ignore evidence from marginalized agents—the model becomes my base model. For this variation, I simulate cases with $m \in (0, 1)$, i.e. I consider cases where dominant agents devalue but do not completely ignore evidence from the marginalized group. The higher the value of m, the more dominant agents devalue.

Because dominant agents still positively update on evidence from marginalized agents, polarization is no longer a stable end state. The two remaining stable end states are community convergence to the true belief and community convergence to the false belief. Thus, for every simulation run, marginalized and dominant groups end with the same belief state. They learn the truth with the same frequency.

4.2 Results

I simulate this model using the following values for the key parameters, for $10,000$ runs each:

- Size of the network (k) : 3, 6, 12, 18.
- Proportion of the marginalized group (d): $\frac{1}{6}$, $\frac{1}{3}$, $\frac{1}{2}$, $\frac{2}{3}$.
- Number of pulls per round (n): 1, 5, 10, 20.
- Probability of B (P_B): .51, .55, .6, .7, .8.
- Degree of devaluation (m): .2, .5, .8.
- Network structure: complete.

I run the model with perfect testimonial reciprocity for all parameter values except for m for comparison.

I find that the marginalized group arrives at the true belief faster than the dominant group (Figure 6).[33] The marginalized group's advantage in learning speed depends on both m and d. As m increases, the difference in the average rounds to successful learning between the two subgroups widens. As d increases, the difference in learning speed also widens. Moreover, marginalized agents select the epistemically better action more frequently than dominant agents during the learning process.[34] As m increases, the difference in frequency between the two subgroups widens.[35]

[33] Measured in the first way introduced in §2.2.

[34] This result is robust for parameters such that the average rounds to successful learning is > 2 for the entire community.

[35] This result is robust for parameters such that the average rounds to successful learning is > 3 for the entire community. The degree to which marginalized agents obtain this epistemic advantage similarly depends on d, but the result is not as robust, especially when m is small.

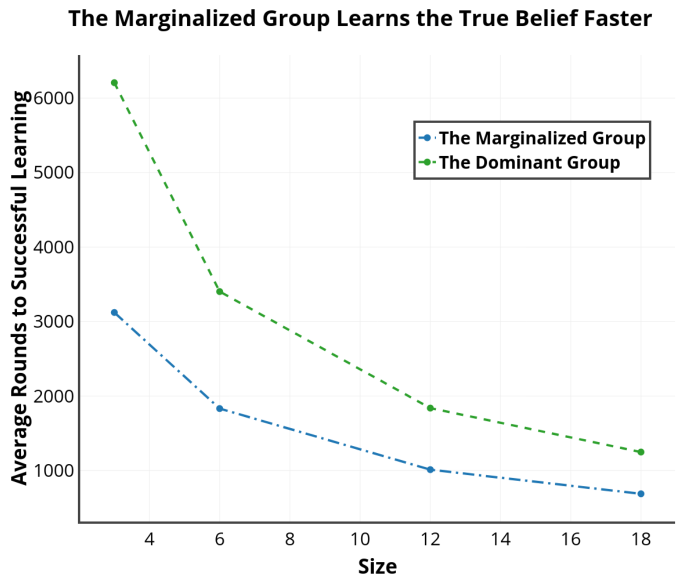

Figure 6: Variation 2, $P_B = .51, n = 1, d = \frac{2}{3}, m = .8.$, $10,000$ simulation runs.

Compared with the model with perfect testimonial reciprocity, the entire community in this variation learns the true belief more slowly.[36] Moreover, as m increases, the entire community in this variation learns more slowly; as d increases, it also learns more slowly. This shows that one-sided testimonial devaluation is detrimental to the entire community's learning speed, and the adverse effect becomes more salient the more dominant agents devalue marginalized agents' testimony.

Finally, comparing the community learning accuracy between this variation and the community with perfect testimonial reciprocity does not bring robust results. The frequency at which the entire community arrives at the true belief fluctuates around that of the model with perfect testimonial reciprocity. In general, as m and d grows, the entire community is more likely to arrive at the true belief more frequently than the community with perfect testimonial reciprocity. What is robust, however, is that the community in this variation always learns the true belief less frequently than the marginalized group in my base model.

5 A Network Standpoint Epistemology

In the above three models, marginalized agents end up with several epistemic advantages, by virtue of their testimony being ignored or devalued by the dominant group. Here, the testimonial

[36] Measured in the second way introduced in §2.2.

ignoration and devaluation practiced by the dominant group is largely epistemically detrimental to its members and the entire community, but is epistemically advantageous to the marginalized group. The inversion thesis, which states that marginalized social groups hold epistemic advantages, is a key claim of standpoint epistemology, though its interpretations and justifications are often contested (see Intemann, 2010; Toole, 2020; Wylie, 2003). My modeling results contribute to standpoint epistemology in two ways. First, I provide a clear interpretation of the inversion thesis by making epistemic advantages precise using several measures. Second, I provide one possible way in which the inversion thesis can arise by showing that it follows from another key claim of standpoint epistemology, namely, the unidirectional failure of testimonial reciprocity.

Standpoint epistemology started as an application of Marx's analysis of the proletarian standpoint to the effect of the sexual division of labor in knowledge production (Hartsock, 1983). It was later extended to cover other unequal power relations' influence on knowledge production. For Hartsock (1983, 298), women's material lived experiences, such as their "relationally defined existence, bodily experience of boundary challenges, and activity of transforming both physical objects and human beings," foster more accurate beliefs for all human activities.

Hartsock (1983)'s argument faces a number of interpretive and justificatory questions. For one, it is unclear exactly what she means by "more accurate beliefs," or more broadly, epistemic advantages. Moreover, it is unclear whether the lived experiences she cites are descriptively true for all women, and it is unclear how they further lead to epistemic advantages. Lurking in the background is also a question about intersectionality—as individual human beings are subjected to different dimensions of oppression, how do we identify the subgroups that hold the most superior knowledge (see Longino, 1993)? Do we look to the ones who are the "most" oppressed within the oppressed groups for the "best" knowledge? These interpretive and justificatory questions have led to heated debates.[37] For instance, Harsock's argument is charged by Hekman (1997) with essentializing women, though Hartsock vehemently denies the charge (Hartsock, 1997). Partly due to intense debates and despite its fruitful applications, standpoint epistemology has been marginalized in contemporary philosophy (Toole, 2020).

This paper is part of a recent effort (e.g. Saint-Croix, 2020; Toole, 2020) at addressing issues

[37] For instance, Alison Wylie (2003) calls standpoint theory "one of the most controversial theories to have been proposed and debated in the 25-30 year history of second wave feminist thinking about knowledge and science."

facing standpoint epistemology by articulating interpretations of the theory that are neither obviously false nor trivially true,[38] and offering explanations for its claims using novel philosophical methods. First, I precisely interpret a subgroup's epistemic advantages using three measures in my models—the frequency at which a subgroup eventually learns the true belief, the speed at which a subgroup learns the true belief, and the frequency at which a subgroup selects the epistemically better action during the learning process. Second, my modeling results show one possible way in which the inversion thesis can be true; namely, when there is unidirectional failure of testimonial reciprocity. Thus, I provide a sufficient condition for the inversion thesis under reasonable assumptions, but not a necessary one. One might regard the claim that marginalized groups' testimony is ignored or devalued as far less controversial than the inversion thesis. Insofar as this is right, my models also have the virtue of explaining a controversial thesis by showing that it follows from something more widely accepted.[39] Given the widespread nature of one-sided testimonial ignoration and devaluation, my results may also shed light on real world cases where the inversion thesis holds.

Note that the mechanism I identified—the unidirectional failure of testimonial reciprocity—is not necessarily the ones that standpoint epistemologists such as Hartsock had in mind. Perhaps it will turn out that women's lived experiences, as Hartsock understands it, lead to their testimony being ignored or devalued, or perhaps Hartsock's reasoning would independently lead to the inversion thesis. In these cases her original justification would still be vindicated. Investigating this, however, is beyond the scope of this paper.

My models admittedly have a few limitations. To start, all agents in my models face the same learning problem, thus they share the same "reality." However, many works in philosophy of race[40] concern the fundamentally different realities faced by marginalized and dominant groups, for instance, the race disparity in policing in the US. Some might suggest that the marginalized group has more accurate beliefs because they are "fluent" in both worlds (see Mills, 2007). My models do not incorporate this notion of "dual realities." As such, though I show that the marginalized group has epistemic advantages when learning the shared "reality," I do not rule out the possibility that the marginalized group might have other kinds of (dis-)advantages due to "dual realities." I plan to

[38] Contra Intemann (2010)'s comments.

[39] Thanks to an anonymous reviewer for raising this point.

[40] E.g., the notion of "double consciousness" in Du Bois (2008).

explore models without a shared "reality" in my future work.

Moreover, my models only have two groups, and I do not consider more groups or agents with multiple group membership. One would expect models with more groups to follow similar epistemic trends, but with slightly altered dynamics. For instance, a testimonial ignoration model with a third "bridge" subgroup might not have polarization as an end state.

In addition, one might worry that some of the assumptions in bandit models may be too idealized. For instance, in some real world situations, investigating one hypothesis may bring insights into other hypotheses as well. One would expect that in this situation, marginalized agents would still gain epistemic advantage in learning speed and the frequency of selecting the better action, though the entire community would eventually be able to reach the true belief reliably.[41] Thinking about how some of the assumptions can be relaxed is a worthwhile direction of future research. Moreover, as Wu and O'Connor (2021) recently notes, some of the network effects in the bandit model paradigm, such as the Zollman effect, are independently discovered in other modeling paradigms like the NK landscape model (Fang et al., 2010; Lazer & Friedman, 2007). It would be worthwhile to test if the marginalized group would end up with epistemic advantages as we apply one-sided testimonial ignoration and devaluation to the NK landscape model. If the results replicate, then this could indicate that the ways in which these modeling paradigms differ are in some sense irrelevant to the phenomena that we aim to explain (see Batterman and Rice (2014)).

Furthermore, in my models, the marginalized agents are participating members of the epistemic community in the sense that they still have access to others' evidence. However, in real epistemic situations, sometimes the very manifestation of marginalization is the *exclusion* of certain agents from epistemic communities. This concern would rightly constrain the applicability of my models. But I would like to suggest that this reflects a merit of my approach as well. Recall that Hartsock's argument faces the following problem of intersectionality: if marginalized groups have epistemic advantages for all human activities, then should we go to the most marginalized group in the world to seek the best knowledge? This problem becomes more puzzling given Narayan's observation that "oppression is often partly constituted by the oppressed being denied access to education and hence to the means of theory production" (Narayan, 1988, 36). My models resist this slippery slope by

[41] This is because all agents always get information about both actions in some form.

focusing on situations where marginalized agents are participating members of the epistemic communities, in that they have access to dominant agents' evidence. In fact, when dominant agents refuse to share their evidence with the marginalized group, the two subgroups effectively function as isolated communities. In this case, the marginalized group, often also in the minority, may learn worse as a result.[42] Moreover, the variation I presented in §3 with homophilic networks, where p_{outgroup} is small, can be reinterpreted to model a version of epistemic exclusion. Here, marginalized agents have very sparse access to dominant groups' evidence, but their informational access to the dominant group is not completely cut off. This is akin to the all-to-familiar situation where very few marginalized knowers are invited to participate in dominant epistemic spaces, but their testimony is still ignored (see, e.g., Settles et al. (2020)). My results from §3 suggest that marginalized agents in this situation still learn the true belief more frequently.

Finally, I will preempt a tempting but misguided response to my modeling results. One might suggest that, since the marginalized group ends up with epistemic advantages, we should now start to ignore or devalue testimony arising from some members of our community, as long as we eventually listen to what they say. This response is misguided for two reasons. First, one-sided testimonial ignoration and devaluation is a textbook case of testimonial injustice according to Fricker (2007). It is unjust because the audience is not giving enough credit to the speaker as they justly deserve. Ignoring or devaluing testimony is committing injustice. Second, one-sided testimonial ignoration is epistemically detrimental to the entire community as it learns the true belief less frequently and more slowly. Moreover, in the case of one-sided testimonial devaluation, where the community *might* learn better than the community with perfect testimonial reciprocity, my results show that the marginalized group can always learn better by refusing to share their evidence with the dominant group.[43] When marginalized agents are in a situation where their evidence is constantly devalued, they would have little incentive to continue sharing their epistemically better informed evidence with dominant agents. Consequently, the entire community would be epistemically worse off, since the learning situation reverts to the base model. Rather than

[42] To be sure, non-participating marginalized agents could have epistemic advantages in other aspects than the shared "reality" for all agents; Narayan (1988) offers a few examples of these. Moreover, when dominant agents refuse to share their evidence with marginalized agents, but still updates on evidence shared by the marginalized group—an epistemic exploitation scenario that may underlie some real cases—, my base model can be reinterpreted to account for this situation too, with the dominant and marginalized groups, and thus their epistemic (dis-)advantages, exchanged.

[43] My base model applies to both when dominant agents do not update and when marginalized agents refuse to share.

prompting individuals to ignore or devalue certain community members' evidence, I hope my modeling results would motivate individuals and communities to identify cases of preexisting failure of testimonial reciprocity, to give epistemic credit where it is overdue, and to recognize the epistemic advantages that marginalized agents may already hold.

Besides offering one possible way in which the inversion thesis could arise by casting it as a consequence of the unidirectional failure of testimonial reciprocity, my modeling results complicate the understanding and applications of certain network effects. Before closing the paper, I will briefly discuss my contribution to network epistemology and note directions of future work.

As previously mentioned, the Zollman effect is usually understood as a claim that "a sparser network structure can benefit an epistemic community" (Rosenstock et al., 2017). Zollman (2007, 2010) finds that the more connected the networks, the less frequently but faster the community learns the true belief. However, Rosenstock et al. (2017) tests Zollman (2007, 2010)'s models using an expanded range of parameter values, and finds that the Zollman effect is not robust for a considerable portion of the parameter space. It is worth noting that Rosenstock et al. (2017) does not find a reversal of the Zollman effect, i.e. more sparsely connected communities never learn the truth less frequently than the more connected counterparts.

The results of my base model further complicate our understanding of the Zollman effect. The community in my model learns the true belief less frequently *and* more slowly than the community with perfect testimonial reciprocity. This shows that for the benefits of the Zollman effect to obtain, the sparse structure cannot manifest in a cutoff of information channels for a subgroup. Otherwise, the Zollman effect may be reversed. For the marginalized group in my base model, on the surface level its members gain epistemic benefits because they do not lose connections, so it seems to counter the spirit of the Zollman effect. However, the marginalized group gains epistemic benefits precisely because they get information from another group that is disconnected. In a sense, the reasoning behind the Zollman effect explains the situation here: the marginalized group benefits from the disconnectedness of the dominant group, and as a result, its members learn the truth more frequently.

This also constrains how the Zollman effect could be applied to real epistemic communities. If, say, a group of scientists decides to interpret the Zollman effect as suggesting that they should stop reading papers from others in the scientific community, but they continue to publish and post to the arXiv, then my modeling results show that as long as the authors of the papers that they ignore continue to read their papers, members of this group may learn worse according to all measures behaving that way.

Earlier I mentioned another interpretation of my base model, which treats the loss of testimonial reciprocity not as the audience refusing to update, but as the speaker refusing to share. Under this interpretation, the audience in general no longer commits testimonial injustice. As it turns out, this alternative interpretation has fruitful applications in social epistemology. For one, it provides another instance of the Independence Thesis, which roughly states that the prescriptions for individual and group decision-making can come apart (Mayo-Wilson et al., 2011). Indeed, a subgroup may learn the truth more frequently by refusing to share their evidence with outgroup members; but the entire community suffers as a result. Industry scientists who have proprietary knowledge but still have access to information from academia can be modeled this way. This interpretation also applies to situations where the dominant group makes their evidence inaccessible to marginalized groups, but exploits evidence from marginalized perspectives. In both cases, the industry scientists and the dominant group would have epistemic advantages. Moreover, applied to scientific communities, this interpretation sheds light on the recent debate on whether the communist norm, which prescribes that scientists share their findings as widely as possible, is an additional contract that scientists should sign (Heesen, 2017; Strevens, 2017). My results would suggest that by following the communist norm, scientists may not learn the true belief as frequently as theoretically possible, but they avoid the epistemic pitfall when no one shares. I explore this interpretation in a follow-up paper (Wu, 2022).

Acknowledgements

I am grateful to Liam Kofi Bright, Adam Chin, Carolina Flores, Nathan Gabriel, Carole Lee, Helen Meskhidze, David Mwakima, Cailin O'Connor, Davin Phoenix, Hannah Rubin, Laura Ruetsche, Jim Weatherall, Alison Wylie, Kino Zhao, and Kevin Zollman for their invaluable comments and helpful discussions on this project. Thanks to members of the Networking for Modelers research group at UCI, audience members at Social Dynamics at UCI in Fall 2020, especially Louis Narens and Brian Skyrms, participants in the Race and Justice Writing Seminar at UCI in Spring 2021, my LPS colleagues at my WiP talk in Fall 2021, and audience members at the PSA 2021 for helpful discussions on the topic. Thanks to three anonymous reviewers for their kind encouragements. Thanks David for support, encouragement, helpful discussions, and thoroughly copy-editing the paper. This paper was made possible in part through the support of grant #1922424 from the National Science Foundation. Simulation codes used in this paper are available on https://jingyiwu.org.

References

Bala, V., & Goyal, S. (1998). Learning from neighbours. *The review of economic studies*, *65*(3), 595–621. https://doi.org/10.1111/1467-937X.00059

Baldwin, J. (1993). *Nobody knows my name: More notes of a native son*. Vintage.

Batterman, R. W., & Rice, C. C. (2014). Minimal model explanations. *Philosophy of Science*, *81*(3), 349–376. https://doi.org/10.1086/676677

Bokulich, A. (2014). How the tiger bush got its stripes: 'how possibly' vs. 'how actually' model explanations. *The Monist*, *97*(3), 321–338. https://doi.org/10.5840/monist201497321

Bright, L. K. (forthcoming). Ida b. wells-barnett's the red record. In E. Schliesser (Ed.), *Neglected classics of philosophy, ii*. Oxford University Press.

Collins, P. H. (2002). *Black feminist thought: Knowledge, consciousness, and the politics of empowerment*. routledge.

Deo, M. E. (2019). *Unequal profession: Race and gender in legal academia*. Stanford University Press.

Dotson, K. (2011). Tracking epistemic violence, tracking practices of silencing. *Hypatia*, *26*(2), 236–257. https://doi.org/10.1111/j.1527-2001.2011.01177.x

Du Bois, W. E. B. (2008). *The souls of black folk*. Oxford University Press.

Fang, C., Lee, J., & Schilling, M. A. (2010). Balancing exploration and exploitation through structural design: The isolation of subgroups and organizational learning. *Organization Science*, *21*(3), 625–642. https://doi.org/10.1287/orsc.1090.0468

Fatima, S. (2017). On the edge of knowing: Microaggression and epistemic uncertainty as a woman of color. In K. Cole & H. Hassel (Eds.), *Surviving sexism in academia: Feminist strategies for leadership* (pp. 147–157). Routledge.

Fricker, M. (2007). *Epistemic injustice: Power and the ethics of knowing*. Oxford University Press.

Golub, B., & Jackson, M. O. (2012). How homophily affects the speed of learning and best-response dynamics. *The Quarterly Journal of Economics*, *127*(3), 1287–1338. https://doi.org/10.1093/qje/qjs021

Hartsock, N. C. (1983). The feminist standpoint: Developing the ground for a specifically feminist historical materialism. *Discovering reality* (pp. 283–310). Springer.

Hartsock, N. C. (1997). Comment on hekman's "truth and method: Feminist standpoint theory revisited": Truth or justice? *Signs*, 367–374. https://doi.org/10.1086/495161

Heesen, R. (2017). Communism and the incentive to share in science. *Philosophy of Science*, *84*(4), 698–716. https://doi.org/10.1086/693875

Hekman, S. (1997). Truth and method: Feminist standpoint theory revisited. *Signs: Journal of Women in Culture and Society*, *22*(2), 341–365. https://doi.org/10.1086/495159

Hornsby, J. (1995). Disempowered speech. *Philosophical topics*, *23*(2), 127–147. https://doi.org/10.5840/philtopics199523211

Intemann, K. (2010). 25 years of feminist empiricism and standpoint theory: Where are we now? *Hypatia*, *25*(4), 778–796. https://doi.org/10.1111/j.1527-2001.2010.01138.x

Kant, I. (1960). *Observations on the feeling of the beautiful and sublime* (Vol. 97). University of California Press.

Lange, M. (2000). Is jeffrey conditionalization defective by virtue of being non-commutative? remarks on the sameness of sensory experiences. *Synthese*, 393–403. https://doi.org/10.1023/A:1005207627094

Lazer, D., & Friedman, A. (2007). The network structure of exploration and exploitation. *Administrative science quarterly*, *52*(4), 667–694. https://doi.org/10.2189/asqu.52.4.667

Longino, H. (1993). Subjects, power, and knowledge: Description and prescription in feminist philosophies of science. *Feminist Epistemologies*, 101–120.

Mayo-Wilson, C., Zollman, K. J., & Danks, D. (2011). The independence thesis: When individual and social epistemology diverge. *Philosophy of Science*, *78*(4), 653–677. https://doi.org/10.1086/661777

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual review of sociology*, *27*(1), 415–444. https://doi.org/10.1146/annurev.soc.27.1.415

Mills, C. (2007). White ignorance. In S. Sullivan & N. Tuana (Eds.), *Race and epistemologies of ignorance* (pp. 11–38). Albany, NY: State Univ of New York Pr.

Narayan, U. (1988). Working together across difference: Some considerations on emotions and political practice. *Hypatia*, 3(2), 31–48. https://doi.org/10.1111/j.1527-2001.1988.tb00067.x

Neely, B. (1993). *Blanche on the lam*. Penguin Group USA.

O'Connor, C., & Weatherall, J. O. (2018). Scientific polarization. *European Journal for Philosophy of Science*, *8*(3), 855–875. https://doi.org/10.1007/s13194-018-0213-9

Rosenstock, S., Bruner, J., & O'Connor, C. (2017). In epistemic networks, is less really more? *Philosophy of Science*, *84*(2), 234–252. https://doi.org/10.1086/690717

Rubin, H., & O'Connor, C. (2018). Discrimination and collaboration in science. *Philosophy of Science*, *85*(3), 380–402. https://doi.org/10.1086/697744

Saint-Croix, C. (2020). Privilege and position. *Res Philosophica*, *97*(4), 489–524. https://doi.org/10.11612/resphil.1953

Settles, I. H., Jones, M. K., Buchanan, N. T., & Dotson, K. (2020). Epistemic exclusion: Scholar(ly) devaluation that marginalizes faculty of color. *Journal of Diversity in Higher Education*. https://doi.org/https://doi.org/10.1037/dhe0000174

Strevens, M. (2017). Scientific sharing: Communism and the social contract. *Scientific collaboration and collective knowledge*, 1–50.

Toole, B. (2020). Demarginalizing standpoint epistemology. *Episteme*, 1–19. https://doi.org/10.1017/epi.2020.8

Wells-Barnett, I. B. (2012). *The red record tabulated statistics and alleged causes of lynching in the united states*. Tredition.

Wu, J. (2022). The epistemology of the communist norm in science.

Wu, J., & O'Connor, C. (2021). How should we promote transient diversity in science? http://philsci-archive.pitt.edu/19428/

Wylie, A. (2003). Why standpoint matters. In R. Figueroa & S. G. Harding (Eds.), *Science and other cultures: Issues in philosophies of science and technology* (pp. 26–48). Routledge.

Zollman, K. J. (2007). The communication structure of epistemic communities. *Philosophy of science*, *74*(5), 574–587. https://doi.org/10.1086/525605

Zollman, K. J. (2010). The epistemic benefit of transient diversity. *Erkenntnis*, *72*(1), 17–35. https://doi.org/10.1007/s10670-009-9194-6