

The Epistemic Significance of Social Pressure

Hrishikesh Joshi

penultimate draft; forthcoming in *Canadian Journal of Philosophy*

ABSTRACT: This paper argues for the existence of a certain type of defeater for one's belief that P—the presence of social incentives not to share evidence against P. Such pressure makes it relatively likely that there is unpossessed evidence that would provide defeaters for P. For, it makes it likely that the evidence we have is a lopsided subset. This offers, I suggest, a rational reconstruction of a core strand of argument in Mill's *On Liberty*. A consequence of the argument is that on morally and politically laden issues in particular, a high degree of doxastic openness might be appropriate.

Keywords: social epistemology, evidence, defeaters, skepticism, John Stuart Mill

1. Introduction

We are deeply social creatures. Not only do we inevitably rely on others for our material and psychological needs, we also rely on others for most of what we know (cf. Sloman and Fernbach 2017). In general, a large portion of our beliefs concern matters that we cannot verify ourselves. I believe that there is a country called Brazil, that Neptune is one of the planets, and that there are electrons, even if I have never been to Brazil or seen Neptune or detected electrons (except in very indirect, “theory laden,” ways like staring at a computer screen).

Traditionally, philosophers like Descartes, Locke, and others focused on the *individual* when they investigated epistemological questions. In recent times, however, there has been more attention given to the epistemic implications of the fact that we are socially situated knowers. Goldman's (1999) book has served as an important jumping off point for contemporary

discussion. However, the core insight that our knowledge rests on social norms and institutions is also prominent in John Stuart Mill (2008) and W.K. Clifford (1877).

Being epistemically interconnected in this way is an enormous blessing, because if we had to verify everything by ourselves, we would not get very far. Nonetheless, this interdependence also carries risks. False or unjustified beliefs, when they affect others—especially those whom we trust—can transfer over to us, and vice versa. If we are to come to an accurate picture of the world then, the social structures underlying the discovery and dissemination of information have to be set up right. Policy makers sometimes talk of proposals needing to be “incentive compatible.” The main idea is that a policy needs to be able to be followed by heterogeneous agents with their own distinct goals. To the extent that a sufficient number of people have an incentive not to behave in a way required by the aims of the policy, it is unlikely to succeed.

It is then an interesting and important question what kinds of social institutions and norms are optimal for the epistemic goals we might endorse. In this vein, Buchanan (2004) has argued that liberal institutions are generally better suited for securing important epistemic goods than non-liberal ones. Addressing the case of modern scientific practice in particular, Longino (1990) argues that a field of inquiry can secure objectivity only if it is robustly open to “transformative critique.” Nguyen (2018) gives a characterization of “echo chambers,” and discusses their deleterious consequences and potential interventions. Mill (2008, 24) makes the striking claim that only the “complete liberty of contradicting and disproving” a claim can properly undergird our having high confidence in it.

In this paper, I want to defend a claim in the vicinity of Mill's, in part by employing Ballantyne's (2015) analysis of *unpossessed evidence* and its epistemic upshots. Roughly, the argument is this. When people face a sufficiently strong incentive not to share evidence for a particular claim P, they will act in line with this incentive. But this reluctance is bad for our collective epistemic health. For it makes it likely that we have only a biased, or unfair, sample of the total evidence out there. Thus, it makes it more likely that the evidence we do not possess contains defeaters for our belief in P. This evidence *about* our evidence is itself a defeater for P. In the presence of such incentives (which I will discuss under the heading of "social pressure"), we ought to be more "doxastically open" (cf. Ballantyne 2019) with respect to P. Further, I will argue that this analysis makes some of the claims Mill makes in *On Liberty* more plausible than they seem at first blush. The paper is thus partly an attempt at a rational reconstruction of one core strand of argument in *On Liberty*, particularly chapter 2.

2. Social Pressure and Unpossessed Evidence

Consider an individual who forms the belief that P on the basis of various pieces of first-order evidence. Suppose that relative to this evidence, the belief that P is justified. But now suppose further that she lives in an environment where there is social pressure to avoid giving evidence *against* P. This may take various forms—perhaps the act of giving reasons that count in favor of not believing P is socially frowned upon and can strain friendships, or perhaps giving such reasons is professionally detrimental.

With some important qualifications to be added later, I claim that the existence of such social pressure to avoid giving evidence against some claim P is at least a partial, *prima facie*, defeater for the individual's belief that P. Whereas a full defeater is a reason to give up one's belief in P, a partial defeater is a reason to reduce the confidence with which one believes P. Partial defeaters do not call for full suspension of judgment. Furthermore, the defeater is *prima facie* because there may be defeater-defeaters, which can make belief in P reasonable despite the presence of social pressure of the kind discussed here. However, if such defeater-defeaters are unavailable, then upon becoming aware of the existence of such social pressure, the individual ought to reduce her confidence in P to a level that is lower than what is warranted by her first-order evidence taken alone.

The reason why the individual ought to reduce her confidence in P is that the presence of this social pressure makes it relatively likely that there is *unpossessed* evidence that would defeat, partially or fully, her belief that P. In other words, the presence of social pressure makes it more likely that this individual's evidence for P is not a representative sample of the evidence out there which bears on P.¹ This is because we are in a deep sense epistemically interdependent. As John Hardwig (1991, 693) notes, "modern knowers cannot be independent and self-reliant, not even in their own fields of specialization." We rely upon others for much of the evidence we have, and so, in order for our evidence for a particular claim to be representative, the social processes which deliver that evidence to us must be structured properly. As socially situated knowers, our

¹ Ballantyne (2019b, 188) characterizes this notion in the following way: "a sample of evidence represents a body of evidence with respect to a proposition *p* only if the sample captures *enough* of the body's evidential features, so that the sample makes reasonable the same type of attitude to *p* that the total available evidence makes reasonable."

beliefs are reliably formed only to the extent that the social incentives at large with regards to those beliefs are right.

Social pressure, especially insofar it is applied independently of the quality of the relevant evidence, can act as a filter on the evidence that makes its way to us, thereby making the evidence un-representative. It can make it the case that the evidence which does make its way to us is a lopsided sample of the evidence out there relevant to the claim at hand. Below is a case which brings out how this might occur.

Suppose that three engineers are responsible for the construction and maintenance of a particular dam. Suppose also that the construction of the dam is a big boon for the community, providing much needed power and irrigation. Further, many local officials and politicians are invested in the success of the dam, which is a huge feather in the cap for them. As a result, there is social pressure not to cast doubt on the project or its success.

Now, imagine there is some good evidence that the dam will break. There are three main reasons to think this, but importantly, these reasons are *distributed* among the engineers. There are also good reasons to think the dam will hold, but these are *shared knowledge* among the engineers. Suppose now that each of these reasons have equal weight and add up in a linear fashion. So, the total evidence suggests the dam is going to break; the three reasons outweigh the two.

To fix ideas, suppose the following are the relevant pieces of evidence:

Evidence that the dam will hold:

E_1 = the dam is made with the proper materials.

E_2 = the structural engineering is well done.

Evidence that the dam will break:

E_3 = the rainfall has been unusually high this year.

E_4 = the spillway design has defects.

E_5 = the outlet pipe maintenance is suboptimal.

Now, suppose the first engineer knows the set $\langle E_1 E_2 E_3 \rangle$, the second knows $\langle E_1 E_2 E_4 \rangle$, and the third knows $\langle E_1 E_2 E_5 \rangle$. If she looks only at the first-order evidence she has, each engineer can rationally conclude that the dam will *not* break. However, due social pressure, each engineer has only a lopsided subset of the total relevant evidence. On the one hand, there is no pressure to avoid sharing evidence supporting the claim that the dam will hold. On the other hand, there is social pressure to avoid sharing evidence suggesting that the dam will break.

Should an engineer in such a situation be confident that the dam will hold? I argue that she shouldn't—or at least, she should be less confident than would be warranted relative to her first order evidence alone. For all she knows, she may be lucky: maybe she has the only piece of evidence that counts in favor of thinking the dam will break. But, for all she knows, the others may be in a situation similar to her, where they have a piece of evidence that they don't share due to social pressure. Even uncertainty regarding whether one's evidence is unrepresentative furnishes (at least partial) defeat for the relevant belief. Once she is aware of the social pressure, she thus ought to appropriately reduce her confidence that the dam will hold, and perhaps suspend judgment.²

² This example is adapted from Joshi (2021).

Empirical work suggests that deliberating groups often tend to focus on evidence that is known by all or most members, resulting in the discounting of unique information as input to deliberation.³ This phenomenon has come to be known within the psychology literature as “hidden profiles.” Social pressure to avoid challenging the dominant opinion within a group only worsens the evidential situation of individuals as they deliberate.

Thus, awareness that there is social pressure to avoid sharing evidence for particular claims undermines the extent to which we can be reasonably confident about those claims. This is because the presence of social pressure makes it more likely that there is unpossessed evidence regarding P such that were it to be added to our first-order evidence, it would defeat our belief that P (cf. Ballantyne 2015; 2019). The presence of social pressure then gives us some important evidence *about* our evidence—and by doing so, undermines the justification for those of our beliefs that are bolstered by it.⁴ That said, depending on the situation, this higher-order evidence need not rationally impel us to reduce confidence in P to the extent that would be required were the defeating first-order evidence to *actually* surface.

Now, in some cases we may nonetheless have good reason to remain confident in P insofar as we have reason to think that, despite the social pressure, the evidence we have is in some sense representative of the evidence out there. In other words, adding all the extra evidence out there to our total evidence would not change much—the total evidence would still point the same way.

³ See, for example: (Stasser and Titus 1985; 2003; Stasser, Abele, and Parsons 2012; Sunstein 2019)

⁴ It very plausible that evidence about our evidence is itself evidence, that it behooves us to incorporate. For a defense of this point, see for example, Feldman (2006). See also Vavova (2018) for a more recent discussion of the epistemic import of higher-order evidence.

Thus, we have a defeater-defeater. However, as I will argue below, such defeater-defeaters are difficult to come by. Because of our epistemic interdependence then, social pressure to avoid giving evidence degrades the confidence with which we ought to hold some of our beliefs.

Consider another simple model for the sake of illustration. Suppose there is an opaque jar that contains two colors of marbles—red and blue. There are 100 marbles in total within the jar. The jar is shaken so as to ensure a randomized mixture. Then, you take a sample of 10 marbles from the top of the jar and end up with 8 blue ones and 2 red ones. Now consider the proposition, $P =$ *there are more blue marbles in the jar than red*. You can be reasonably confident that this is true. Of course, there is a small chance that due to a sheer fluke, you happened to get too many blue marbles. We can't *rule out* this possibility. But given that you took a reasonably sized sample in this case of 10 marbles, you can be justified in placing relatively high confidence in P . And you can be more and more confident as we increase the sample size. Note that in this case there is lots of unpossessed evidence out there—you haven't looked at the remaining 90 marbles. However, this is not too troubling. The reason is that you have a *fair sampling* of the total relevant evidence. The mere fact that there is (lots of) unpossessed evidence does not significantly undermine the confidence you can rationally place in P .⁵

Now let's change the case. Imagine that you get the same result—8 blue marble draws and 2 red draws. However, you learn that red marbles have been made of out of a heavier material than the blue. Heavier marbles tend to sink to the bottom relative to lighter marbles, when jars containing a mixture are shaken. Awareness of this feature of the setup should cause you to significantly

⁵ This example is adapted from Ballantyne (2015).

reduce confidence in P. Depending on the difference in the weights, the length of time the jar is shaken, etc., it might be appropriate to suspend belief in P or even think that not-P is more likely than P (despite the weighting, 2 red marbles made it to the top!).

The analogy with social pressure is straightforward. Social pressure to avoid giving evidence against some proposition P is akin to the red marbles having greater weight. To the extent there is such social pressure with respect to P, we can expect to have an unfairly biased set of first-order evidence as it pertains to whether P. The greater the social pressure, the more unfair our sample of evidence is likely to be. The presence of such social pressure then serves as a *prima facie* defeater for one's belief that P.

The defeater is only *prima facie* because there can be defeater-defeaters. Going back to the urn case, imagine that though the red marbles are weighted, they are weighted with iron filings, and the room in which you are drawing the marbles contains an upward pulling magnetic field. Here, depending on the amount of iron filings per red marble, the strength of the magnetic field, etc., your sample may well be fair, or even biased in the other direction. If you know that the relative effects of these two forces cancel out exactly, for example, then you can hold on to your original confidence in P.

But notice that such additional factors make reasonable belief in P more difficult to secure. For, even uncertainty about how the magnetic pull compares with the additional weighting furnishes defeat for P. In general, if one is uncertain about whether one's evidence regarding P is representative, that uncertainty is epistemically troubling—its presence calls for a different, more

agnostic, doxastic attitude towards P than what the first-order evidence taken by itself would recommend. In a similar vein, we might claim that though the existence of social pressure is a *prima facie* defeater, other factors may render it epistemically harmless (or even beneficial). However, as will be discussed later on, such defeater-defeaters are often difficult to come by.⁶

3. Mill on Social Pressure

The above model of how the existence of social pressure to refrain from providing evidence for certain claims undermines justification helps to make sense of some of John Stuart Mill's striking claims in *On Liberty*. Mill lays out a fairly demanding condition for when we can be justified in believing certain types of claims. That condition is the "complete liberty of contradicting and disproving our opinion" (Mill 2008, 24). Mill goes on to say: "if even the Newtonian philosophy were not permitted to be questioned, mankind could not feel as complete assurance of its truth as they now do" (p. 26).

As he makes clear at the outset in chapter 2 of *On Liberty*, Mill is not talking about governmental or formal restrictions on what can be discussed or disputed. Rather, he is interested primarily in social and professional pressures we might face to avoid disputing certain claims. Indeed, on his analysis, even legal penalties owe much of their efficacy to the social pressure that they can create. "The chief mischief of the legal penalties," he writes, "is that they strengthen social

⁶ The urn example assumes a simple, stylized model of how evidence works. There are situations where collecting data is akin to sampling marbles from the jar—for instance, public opinion polling. In other cases, evidence can work in more complex ways. Nonetheless, evidence that one's first-order evidence is a lopsided subset of the evidence out there is still epistemically troubling as a general matter.

stigma. It is that stigma which is really effective.” This stigma is especially powerful when it has professional consequences—loss of current employment or opportunities for employment, for example. Hence, “men might as well be imprisoned, as excluded from the means of earning their bread” (Mill 2008, 37).

On the face of it, the claim that we can only be reasonably confident in views for which there is minimal social and professional sanction attached to disputing them seems too strong. What ultimately matters, it might be thought, for whether we are justified in believing some claim P is whether the available relevant evidence supports P. Rewind back to 1859, when Mill published *On Liberty*—what matters in whether a physicist is justified in believing Newtonian physics is the evidence available to the scientific community. If that evidence (constituted by observations, scholarly publications, etc.) supports Newtonian physics over rival theories, then belief in that theory is justified. Why should it matter whether it is “permitted to be questioned”?

This response misses the significance of unpossessed evidence. To be justified in believing something, to reasonably place high confidence in it, we need a robust mechanism to ensure that whatever relevant evidence may be out there, which we can (in some sense) have access to, is brought to light. This unpossessed evidence may be of two kinds. First, it might be possessed by individuals but not shared with the community or group at large—this is what happens in the case of the engineers and the dam in the previous section. Alternatively, the evidence might be out there, waiting to be discovered. This, as it turns out, was the case with Newtonian physics in 1859. Albert Einstein would later propose his special and general theories of relativity, and experimental data eventually disconfirmed Newtonian physics in a range of special scenarios.

Thus, one way to make sense of Mill's seemingly high bar for justification is that social and professional pressure can hamper the ability of unpossessed evidence to surface and be widely known in two ways. First, individuals who have a piece of disconfirming evidence might not share it with the wider community—especially if that piece is small enough not to undermine the overall case for a particular claim. Why attract social stigma just to share a contrary piece of evidence that leaves our general impression of the world largely unchanged? Second, the presence of social and especially professional pressure will discourage attempts to disconfirm the widely held view. To return to Mill's example, imagine there was sufficiently strong stigma attached to challenging the assumptions of Newtonian physics. This would discourage potential Einsteins from embarking on research programs that undermine Newtonian physics. Because of these two ways in which social pressure can prevent unpossessed evidence from reaching us, its presence is at least a partial, *prima facie* defeater for our relevant beliefs.

4. Ideal and Non-Ideal Epistemology

Here, I want to examine an objection that might be pressed against the argument sketched above. Addressing it will bring out some of the presuppositions of the Millian view described above. I will suggest that these presuppositions find some support from recent empirical work in psychology and cognitive science.

The objection goes like this. Evidence can be misleading. Moreover, individuals are susceptible to a range of cognitive biases can lead them to over- or under-weigh evidence in a range of

contexts. Perhaps if we all were *ideal epistemic agents*, then the Millian picture of social pressure being an epistemic defeater would make sense. Indeed, in the dam and engineer example used above, it is assumed that the engineers are perfectly rational in responding to their first-order evidence about whether the dam breaks. However, people in general are not perfectly rational in this way. They are prone to biased weighting of evidence, conspiracy theories, and are often easily duped by misleading claims. Given this, social pressure can play a *positive* epistemic role by ensuring that falsehoods and incorrect theories do not receive unjustified uptake.

This is of course not to say that all forms of social pressure to avoid disputing certain views are beneficial to truth-seeking and inquiry. We can acknowledge, for example, that the social, professional, and legal pressures that existed in Galileo's time, on the topic of geo-centrism, were counterproductive. As was the environment that sustained Lysenkoism in the Soviet Union in the 1940s and 50s. But these are special cases. It's not a general rule that social pressure to refrain from expressing certain pieces of (putative) evidence is always counterproductive. When it comes to matters that are settled, and particularly in cases where we have good reason to think that disputing these matters will do more harm than good, social and professional pressure (even if not legal sanctions) are epistemically and morally beneficial.

Now, a plausible Millian view need not be committed to denying that every single case of such social pressure has detrimental consequences. Sometimes the stars can align so as to produce epistemically and morally positive results. But, the Millian will insist, it is a good rule of thumb that social pressure not to dispute certain claims will be detrimental to truth-seeking. As such, its presence undermines the justification we might have in believing those claims. On the model I

have sketched, the presence of such pressure not to provide evidence against P is a partial defeater for the belief that P because it increases the likelihood that there is unpossessed evidence against P that would make belief in P unjustified.

However, even to defend the above as a rule of thumb, the Millian needs to make further assumptions. One key assumption has to be that social pressure typically does not work in a fortuitous way—i.e. that it is not a reliable mechanism of promoting the truth. Second, we are, in general, relatively better at assessing first-order evidence relevant to particular claims than we are at assessing which sorts of social pressure are epistemically beneficial, especially in the long run. These assumptions, are, I think, defensible.

An important clarification is in order. Why care about the long run in this context? One might worry that there is an illegitimate comparison going on here—our synchronic ability to assess first-order evidence is being pitted against our ability to predict the long-term consequences of processes responding to social pressure. In general, it's difficult to predict the long-term consequences of almost anything when it comes to complex social phenomena (cf. Tetlock 2015). While this last point is certainly right, it further brings out the challenges of preserving justification in the presence of social pressures of the sort discussed here, because evidence generation and sharing is almost always a *diachronic* process. I might share a piece of relevant evidence today, you might share something tomorrow, and someone else might discover important evidence next month. To reasonably preserve high confidence in some proposition P despite the presence of social pressure requires us to be in an unusually strong evidentiary

position—namely that of being able to predict that the evidentiary output of a diachronic process of sharing and discovery will not furnish defeat for P.

Now, Mill himself defends the first assumption in part via an inductive argument. When we look at the historical record, we see that in most of the cases where social and legal pressures have been applied against “heretical” opinions, they have not yielded epistemically beneficial outcomes. For instance, he writes that “history teems with instances of truth put down by persecution,” (Mill 2008, 33) and immediately thereafter gives several of what he sees as good examples of this. Moreover, he claims, it is precisely in those circumstances where people feel that particular opinions are not only false but will have pernicious consequences if they spread, and moreover, are “immoral and impious,” where society is likely to apply stigma and sanction in epistemically counterproductive ways. In his words, “these are exactly the occasions on which the men of one generation commit those dreadful mistakes, which excite the astonishment and horror of posterity” (Mill 2008, 29). One way to interpret this argument is that the historical record provides higher-order evidence that under the conditions described above, we are likely to be mistaken. Far from being an exercise in epistemic ideal theory, then, the Millian project can be squarely located as a guide for socially situated, morally and epistemically non-ideal agents. Given the historical record as he sees it, Mill is skeptical of our ability to wield social pressure in epistemically reliable ways.

One bit of support for this Millian skepticism might come from Elisabeth Noelle-Neumann’s (1974; 1984) analysis of social pressure’s role in influencing public opinion in the long run. Her basic model is this. People are, for the most part, afraid of “isolation pressure”—that is, the real

or perceived threat of being isolated or marginalized by others.⁷ Thus, most individuals are on the lookout for whether an opinion is in the ascendancy or decline. When people feel (rightly or wrongly) that expressing a view would make them susceptible to social isolation (particularly from groups important to their life prospects) they tend to refrain from expressing that view. On the other hand, those who see their opinions as in the ascendancy tend to express their views loud and clear. This sets into motion a “spiral of silence” which eventually leads to the former view to become more or less absent from public discourse. These processes typically occur with respect to controversial and morally laden topics. Importantly, the number of partisans on either side does not matter, if the latter, ascendant group is sufficiently vocal and coordinated. The process is also dramatically strengthened to the extent that mass media repeatedly and concordantly side with the ascendant view.

Suppose that this model is roughly accurate. What should we make of its epistemic consequences? The “spiral of silence” is a complex social process, and its trajectory is guided by various social forces—things like: which individuals are louder, who has most influence on mass communication, which groups are most coordinated in promoting their views, etc. But there is no reason to think that these forces robustly track the truth. Or at least, absent some plausible story according to which these forces are robustly truth-tracking, we should expect it to be a fortuitous coincidence if opinions settled through such processes happen to be true. To borrow an example from Sharon Street (2006), suppose you set sail for Bermuda, and allow the ship to be taken

⁷ A classic study supporting this idea is Solomon Asch’s (1955) experiment, where a substantial portion of subjects got a very easy visual exercise wrong simply because they observed others giving the wrong answer. The basic results have been reproduced across several different countries, though there is some variability across cultures (Bond and Smith 1996) and times (Larsen 1990). See also Sunstein (2019) for a recent discussion of the state of the literature and its implications.

wherever the winds blow, without some sort of compass or navigational instrument. You eventually hit land; now how confident should you be that you have landed on Bermuda? Plausibly, not very confident—the direction of the wind has no robust relationship with guiding your boat to Bermuda. If you somehow happen to land on Bermuda, that would be a stroke of luck.

Now, I don't think the Millian needs to be committed to Noelle-Neumann's specific model of the dynamics of public opinion. However, the model is illustrative and the lesson we can draw from it is fairly general. Given that we live in a complex epistemic environment where morally and epistemically non-ideal agents interact strategically with each other, against a backdrop of myriad institutions with their own norms and internal practices, we shouldn't expect social forces—in particular, social pressures to avoid disputing particular claims—to robustly help us converge upon the truth.

The second assumption the Millian has to make is that, in general, our ability to appropriately assess the significance of first-order evidence given to us by others, relevant to some claim P, is better than our ability to judge whether social pressure to avoid giving evidence in support of (or against) P will have positive epistemic consequences in the long run. Note first that the latter type of judgment will require extremely difficult calculations and enormous prescience. Given that our evidential landscape is complex in the way described above, it's hard to be rationally confident about which social pressures (if any) will be optimal in the long run. On the other hand, assessing first-order evidence given to us by others plausibly requires less. However, it

does require us to be epistemically *vigilant* so as to avoid being easily duped by people giving *bad* arguments.

Some recent work in cognitive science suggests that our abilities for epistemic vigilance are more developed and effective than is typically assumed (Sperber et al. 2010; Mercier 2020). The core idea is this. Communication involves a strategic element—it is not always in the interest of the sender of a signal to convey accurate information. For instance, it would be in the interest of a bird to signal to rival birds that a predator is near, even when this is not so, as a means for diverting other birds from a newly found source of food. It would be in the interest of a door-to-door salesman to convince you that his product is useful for your purposes even if it isn't.⁸

Examples like this can be multiplied indefinitely. The crucial point is that for a species like us to have evolved mechanisms for communication, *recipients* must have reliable capacities for vigilance—completely trusting and naive recipients would not have been very successful in producing surviving offspring, for obvious reasons.

5. Varying the Assumptions of Rationality: A Model

Our capacity for vigilance doesn't mean we are free of cognitive biases. Indeed, a large literature in psychology highlights ways in which we can be biased reasoners.⁹ When it comes to evaluating evidence in particular, there are two issues going on. First, when people agree with an argument's conclusion, they tend to evaluate it relatively superficially. On the other hand, when

⁸ Repeated interaction can align some of these incentives (Axelrod 1984). A salesman who conducts business with you repeatedly probably doesn't maximize his long run profits by duping you. Yet even this requires some vigilance on your part, lest he fool you twice. Moreover, not all games are repeated in this way.

⁹ For a useful overview, see Mercier and Sperber (2017).

people disagree with a conclusion, they evaluate the argument more thoroughly and rigorously. This feature of our psychology is sometimes called the “myside” or “confirmation” bias. Second, we are typically more exacting and adept at assessing *others’* reasons than we are at assessing *our own*, even bracketing the content of those reasons (Trousseau et al. 2016). Supposing these are indeed biases that depart from rational evaluation of evidence, should we be more or less sanguine about the epistemic consequences of social pressure?¹⁰

It’s not obvious that we should be more sanguine. For illustration, consider three individuals with different attitudes toward some proposition P:

Arturo doesn’t have a settled opinion as to whether P. [Suppose for simplicity, that he suspends judgment as to whether P.]

Benazir has evidence E that she thinks supports P.

Carol firmly believes that P is false, and thinks that putative evidence for P is either misleading, weak, or just not evidence at all.

Suppose now that Carol is willing and able to wield social/professional pressure against Benazir so as to incentivize her not to share E with Arturo. Should we expect this to be epistemically beneficial or detrimental to Arturo?

Assuming that we are susceptible to “myside” bias, we should be generally skeptical of Carol’s ability to *decide for others* (Arturo in this case) what the evidence supports. Given that she is antecedently committed to P, Carol is likely to evaluate the arguments in favor of P with less scrutiny—the same will not apply to Arturo. Since he has no dog in the fight, we should be

¹⁰ There is some room for rational reconstruction of “confirmation” bias in Bayesian terms.

relatively optimistic about Arturo's ability to take the appropriate attitude toward P after Benazir presents her evidence, E. At least, in general, we should be *more* optimistic about Arturo's ability to draw the proper lesson from E than Carol's ability to determine whether E is best left unspoken.

This holds even as we make the agents more and more ideal. If Carol suffers from no cognitive bias with respect to P, and has been perfectly rational in forming and holding that belief, she might still be unaware of what other evidence *Arturo* has. Thus, suppose that Carol has total evidence E_C and Arturo has total evidence E_A . It might be the case that $E_A + E$, taken in sum, justifies belief in P whereas $E_C + E$ does not. Unless Carol has perfect knowledge of the relevant contents of Arturo's mind, she is unable to rule this out—no matter how careful she has been in holding her own belief that P.

Though he doesn't put things in these terms, Mill appears to have something like this in mind when he presents his case against censorship. The key for Mill is that we shouldn't assume ourselves to be "infallible." For him, this doesn't mean avoiding having any firm beliefs, held with high confidence. That claim would be implausible and would prove too much. On Mill's picture, Carol does not presume herself to be infallible when she places high confidence on P being false. Rather, she assumes her own infallibility when she moves to decide *for Arturo* that he should not hear E (which, on Carol's view is misleading) from Benazir. Mill writes, "it is not the feeling sure of a doctrine (be it what it may) which I call the assumption of infallibility. It is the undertaking to decide that question *for others*, without allowing them to hear what can be said on the contrary side" (Mill 2008, 28). Mill further emphasizes that it is individuals like

Arturo who are likely to be the (epistemic) beneficiaries of Benazir's publicly sharing E. Indeed, Benazir's sharing her evidence may well cause Carol to dig in her heels, depending on how affectively invested she is in the conclusion. But, Mill writes, "it is not on the impassioned partisan, it is on the calmer and more disinterested bystander, that this collision of opinions works its salutary effect" (Mill 2008, 58).

Thus, to the extent that individuals like Carol have the power to incentivize individuals like Benazir to not share evidence (which the former see as misleading), the epistemic position of *others* thereby deteriorates with respect to that subject. In the example above, Arturo has no opinion about P. But imagine another character, David, who though he himself does not seek to wield social pressure like Carol does, nonetheless is fairly confident that P is false, given his—let's assume rational—take on his first-order evidence. David too, should become less confident about P (or, alternatively, more doxastically open) if and when he learns that there is social pressure on individuals like Benazir to avoid sharing evidence for P. This is because such pressure makes it likely that his evidence is not a representative sample of the total evidence out there. This evidence of potential defeaters present in the *unpossessed* evidence is itself a *prima facie* defeater for David's belief that P.

An important qualification is in order. The social pressure that Carol applies can be epistemically beneficial if we assume the following: there is a significant enough gap between the quality and scope of Carol's evidence and her ability to draw the appropriate conclusion from such evidence, and the corresponding evidential sets and relevant abilities of characters like Arturo and David.

But this is precisely what we should expect if Carol is a (genuine) *expert* with respect to P and Arturo and David are not. The implications of this point are explored in the next section.

6. Expertise and Preemption

We live in a time marked by intense division of cognitive labor. The best we can do as individuals is to become experts in one narrow subfield. And often, we lack not only knowledge of the relevant facts, but also facility with the tools needed to make appropriate independent judgments about propositions regarding subject matters that are sufficiently distant from our areas of expertise. Thus, epistemic responsibility calls for appropriate deference to the experts in such cases—it would be epistemically irresponsible to “do your own research” in certain scenarios, as opposed to simply relying on expert judgment.

Moreover, in cases where the tools, conceptual resources, and relevant facts constitute a landscape that is too complex for a novice to navigate by themselves, an expert might rightly *preempt* contrary evidence that is misleading (Begby 2021). The basic feature of such cases is this: the expert possesses enough evidence, as well as proper grasp of the relevant tools for assessing the evidence, to see that some proposition Q, though it counts against (or might look to the novice like it counts against) P, is misleading. Given the total available evidence, the case for P is very strong—nonetheless, novices who do not have the requisite grasp of that total available evidence are liable to interpret Q as refuting P or seriously casting it into doubt. In such a situation, the expert might preempt this by saying for example, “you might encounter Q, but that should not lead you to doubt P.”

Here are two examples from Begby (2021) that bring this out vividly. A study is released with the conclusion that a particular vaccine is associated with increased risk of autism. The expert might know, for example, that the study used a small sample size and that well-done meta-analyses do not establish a link between the vaccine and autism. A novice, however, who might not be aware of cutting-edge statistical methodology in medicine might see the study reported, and through no fault of his own, come to believe that the vaccine causes autism. The expert can preempt this by warning the novice of the study and telling him, without necessarily delving into the complicated reasons (for considerations of time or audience sophistication, say) why the study does not establish the causal link. By means of such preemption, the expert lets the novice know that the evidence provided in the study has already been taken into account. Similarly, consider the climate change skeptical argument which proposes the claim that there is increased solar activity in recent years as a reason to doubt the significance of human industrial activity in causing global temperatures to rise. Again, an expert might preempt this by assuring novices that fluctuations in solar activity are already incorporated within the latest climate models.

Now imagine a novice who, despite being warned by credible experts in this way, nonetheless believes that the vaccine causes autism or that human industrial activity is not causing climate change. What is more, he shares the contrary arguments (which do contain evidence, albeit weak and misleading) with other impressionable novices. We can now imagine people placing some social pressure on this person not to share these arguments—for, not all impressionable novices might be within the reach of expert evidential preemption. It is very plausible that such social pressure is *epistemically beneficial*. This is an important qualification to the thesis of this

paper—in such circumscribed contexts where the evidentiary landscape is sufficiently complex for novices and where credible and appropriately motivated experts have preempted contrary misleading evidence, social pressure can have *positive* epistemic upshots.

However, the core Millian insight can be preserved by noting that this observation in effect pushes the analysis up one level. (It is also noteworthy that Mill himself chooses the case of Newtonian physics, regarding which there would have been easily identifiable experts in his time.) For us to be able to trust experts in this way, and to be rationally guided by their preemption, we must have some assurance that *their* evidential landscape has not been degraded by social pressure. The mere fact that they are credentialed experts does not guarantee that the evidence they possess is representative, if there are counterproductive social pressures operating within their own social and professional spheres. In the dam and engineers case presented earlier in the paper, the engineers are assumed to be experts in the sense that they have the relevant training and experience for dam maintenance. Nonetheless, each engineer has a lopsided subset of the total available evidence—preemption from experts in such situations is of little help.

Begby helpfully characterizes a general principle that is operative here. “In standard cases,” he writes, “whatever it is that justifies my epistemic trust in my testifier’s assertion that *p* also justifies my epistemic trust in my testifier’s assessment of the total evidence that bears on *p*” (Begby 2021, 519). Furthermore, “if it were brought to my attention that I have reason *not* to trust my testifier’s assessment of the total evidence that bears on *p*, then I should immediately recognize a reason to modify my confidence in my testimonially acquired belief that *p*” (Begby 2021, 520). Now, experts are but people who can themselves be subject to social and

professional pressures: thus, a system of incentives can be maintained where each (or most) of the experts within a domain possess only an un-representative subset of the total evidence. In such cases the import is somewhat different—it's not quite that one doubts the expert's ability to properly assess the total evidence that the expert *possesses*, but that one reasonably doubts whether that expert's total evidence is representative of the available evidence out there. This doubt is enough to degrade the justification with which one might hold the belief that P simply based on testimony.

What complicates matters further is the fact that many important questions do not fall neatly within a sub-discipline with respect to which reliable experts can easily be found. Rather, such questions require collaboration from individuals operating in a range of distinct disciplines, as well as input from the relevant novices.¹¹ In such situations, an expert with respect to one sub-discipline cannot easily perform the function of helpful preemption without running the risk of epistemic trespassing (Ballantyne 2019a). Moreover, for a *group* of experts working within distinct subdisciplines to properly assess the evidence they possess *as a group*, they must operate in a social environment where the sort of phenomenon occurring in the three engineers and dam case is mitigated—i.e. they must feel free to share their evidence without fear of (severe enough) social or professional sanction.

¹¹ For example, accurately assessing the water quality within a city might require not only analysis by water treatment experts but also input from local residents who can directly observe their tap water and its effects. Some have suggested that this is a lesson to be learned from the Flint water crisis in Michigan (US); see Pauli (2020).

7. Limitations and Practical Upshots

Here is a stark illustration of the chief point made thus far. Suppose you live under a ruthless authoritarian dictatorship, called Utopia. The regime places severe sanctions on individuals who undermine the message it seeks to convey to the populace. One of these messages is that relative to the outside world, Utopia's economy is doing strikingly well, thanks to its wise leadership.

Now there are several specialized epistemic communities in Utopia tasked with scientific and/or social inquiry. Many of these have nothing to do with the central ideology or messaging priorities of the regime. There are meteorologists who record the weather, physicists who study conductive materials, and chemists who investigate new binding agents. A physicist who finds that material XYZ is a better conductor of electricity than material ABC under cold temperatures does not threaten the message of the regime. Neither does the meteorologist who tells you it will be warm tomorrow with a slight chance of showers.

However, there are also economists. It is widely known that in the not so distant past, economists who have suggested that the indicators are lagging or that there are better ways to structure the economy have never been heard from again. Now the economists who are currently well-established regularly point to good evidence (as far as you can tell) that the economy is exceeding all expectations. They present their graphs and models, and let's even stipulate that as far as you know they do not straightforwardly fabricate data. The graphs and data check out given what you can tell.

Should you believe what the Utopian economists tell you? Plausibly, at the very least, you should take their conclusions with a generous serving of salt even if you can't find errors in their presentation. This is because the social conditions which would allow you to trust them are not present. These economists have extremely strong incentives to only tell you one side of the story. On the other hand, you can justifiably place relatively much higher confidence in the conclusions of the Utopian meteorologists, physicists, and chemists. Sure, individuals can make mistakes sometimes, but given that these communities face little or no pressure to present only one side, their errors will be corrected over time.

This is of course an extreme case and I suspect that few, if any, readers will find it plausible that in fact trust in these economists' conclusions is warranted. Notice though that the extreme case does vividly bring out the idea that which conclusions we can trust, based on the first-order evidence presented to us, is a function of the incentives that others have with respect to presenting the evidence. Ideally, we want people to have incentive to adequately present the evidence they possess. As we depart from this ideal, trust becomes less appropriate. Now, departing from this ideal is not an all-or-nothing matter, and there is bound to be a spectrum here. On the one hand we have the economists of Utopia, and on the other hand we have well-functioning epistemic communities with norms that help them robustly track the truth.

As Mill was keen to emphasize though, it is not only authoritarian governments which can create such misaligned incentives that lead to bad epistemic environments. Social and/or professional pressures can also be detrimental in this regard. Thus, we can imagine epistemic communities malfunctioning because there is social and professional pressure not to argue in favor of some

hypothesis H. This pressure need not be “top down” as would be the case within authoritarian regimes. Rather the pressure can take the following form: defending H reduces job prospects, gets you invited to fewer professional conferences or parties, makes it likely that others publicly distance themselves from you, and so on. Such incentives—though not as strong as those in Utopia—are likely to be effective, especially insofar as individuals are driven to achieve social status, gainful employment, and want to avoid “isolation pressure.”

In modern liberal societies, determining where such pressures function in a detrimental way is somewhat complicated by the fact that it is a standard part of academic practice to highly regulate speech and publication, by methods such as rigorous peer review (Simpson 2020). Such methods by design screen out bad or substandard argumentation or evidence-presenting. Furthermore, there is bound to be some social stigma attached to defending hypotheses that are highly unlikely given our total evidence—for example, that the earth is flat or that anthropogenic climate change is not occurring or that living organisms did not evolve through a process of natural and sexual selection.

What we seem to need then, is a suitable qualification: when social pressure is applied to individuals presenting evidence that is independent of the quality of the evidence, this is where our epistemic situation with respect to that hypothesis is likely to deteriorate. In the case of Utopia, the pressure applied is conclusion driven—economists who present evidence suggesting the economy is faltering are never heard from again regardless of whether their methodology is rigorous. To bring matters closer to home, there is thus an important difference, for example, between the case of an author being criticized or denied professional opportunities for publishing

a study with sloppy statistical methodology versus someone facing these costs for publishing a statistically rigorous study that defends a hypothesis that is unpopular within the field. In this vein, Longino (1990) emphasizes the need for diversity in background assumptions within a field of inquiry as a means for enabling transformative critique when appropriate.

In practice, however, it is bound to be difficult to ascertain which type of scenario we are in with respect to particular hypothesis. Though there are easy extreme cases, the terrain is murkier in the middle. Indeed, it seems to me that we would be expecting too much from epistemology if we demand a precise and determinate litmus test. However, I want to speculatively propose one rough heuristic. Hypotheses in which members of an epistemic community have an *affective investment* are more likely to be the locus of the problematic type of social pressure. Now when it comes to particular hypotheses about the physical world, for example, it is rare to find such affective investment. In some sense, nobody cares about which materials are more conductive, the chemical compositions of particular polymers, the reproductive rates of *E. Coli* bacteria, and so on. Affective investment is much more likely with respect to policy adjacent or ideology adjacent (political, moral, religious, etc.) issues. Indeed, in Galileo's time and place, geo-centrism was close to the core tenets of the prevalent religious worldview. Today, we might expect epistemic communities to have other issues that are core to their worldviews: particular hypotheses about the causes and remedies for various social problems, for example.

Perhaps tragically then, our lights get dimmer just as we start to care more about a particular domain. This is so especially in cases where having a firm position on particular issues become closer to our identities. "Convictions are more dangerous enemies of truth than lies," Nietzsche

(1954, 483) says in *Human, All Too Human*, and the discussion so far suggests one way in which this might be true.

Some recent work in social psychology and political epistemology seems to corroborate this point. Several authors have noted that our identities can have a distorting effect on how we process evidence. For example, Dan Kahan and colleagues (2017) find that political partisans, even those relatively good at numerical reasoning, are prone to making basic statistical errors when the presented data suggest a conclusion contrary to their partisan beliefs.¹² Similar results have been found in the case of assessing the logical validity of arguments (Gampa et al. 2019). Su (2022) finds that subjects are comparatively hesitant to update their beliefs in light of even highly reliable information, when doing so challenges their political beliefs. Furthermore, even highly knowledgeable partisans are prone to these biases (Hannon 2022).

If the argument developed in this paper is right, however, the problem goes deeper than that. Our epistemic predicament in such domains is not merely a function of our individual propensities for motivated irrationality. Rather, because we depend on others for our epistemic health, and because chiming in on morally and politically laden topics is more likely to be subject to social pressures, the evidence we are presented is likely to be a lopsided subset of what might be out there. The individual's mitigating her own biases when it comes to assessing first-order evidence can only get her so far.

¹² There is some controversy over whether the results in the original study replicate. Ballarini and Sloman (2017) claim that the results do not replicate with their independent sample. However, Kahan and Peters (2017) respond that this failed replication attempt is inconclusive because it uses a small (N=55) and ideologically homogenous sample of college students, and provide their own replication with a larger, more diverse sample. However, see Persson et al. (2021) for more recent worries, particularly that motivated numeracy occurs only at the top end of the numeracy spectrum.

8. Conclusion

It is a truism that our perspective on the world is heavily influenced by the social context in which we find ourselves. This is the epistemic dimension of the cooperation that human society is built on. However, just as cooperation in general can break down when the incentives and norms are not right, so can our picture of the world can deteriorate (in accuracy) when the incentives to share evidence are compromised. This is one of the core themes and insights in Mill's *On Liberty*, chapter 2, which is as much an essay in social epistemology as it is a defense of freedom of speech. On the analysis presented here, social pressure to avoid sharing evidence against a particular claim undermines the confidence we can place in that claim, because it makes more likely the possibility that the (first-order) evidence that does make its way to us is a lopsided subset of the total. This has the perhaps tragic implication that we can typically be less confident of morally and politically laden issues than we can about “dry” subjects like chemistry or cell biology. This is a further reason why there might be rational pressure to be doxastically open on such issues despite the natural temptation to have firm opinions.

Hrishikesh Joshi
Bowling Green State University
joshih@bgsu.edu

Acknowledgments: I want to specially thank the two anonymous referees and the handling editor at *Canadian Journal of Philosophy* for extremely helpful and detailed feedback. Thanks also to Sahar Akhtar, Daniel Jacobson, JP Messina, Kranti Saran, Noel Swanson, Justin Tosi, Craig Warmke and audiences/workshop participants at Ashoka University, Georgetown University, Northern Illinois University, University of Delaware, and University of Colorado—Boulder.

References

- Asch, Solomon E. 1955. "Opinions and Social Pressure." *Scientific American* 193 (5): 31–35. <https://doi.org/10.1038/scientificamerican1155-31>.
- Axelrod, Robert. 1984. *The Evolution of Cooperation*. New York: Basic Books.
- Ballantyne, Nathan. 2015. "The Significance of Unpossessed Evidence." *The Philosophical Quarterly* 65 (260): 315–35.
- . 2019a. "Epistemic Trespassing." *Mind* 128 (510): 367–95.
- . 2019b. *Knowing Our Limits*. New York: Oxford University Press.
- Ballarini, Christina, and Steven A Sloman. 2017. "Reasons and the 'Motivated Numeracy Effect.'" *Proceedings of the 39th Annual Meeting of the Cognitive Science Society*, 1580–85.
- Begby, Endre. 2021. "Evidential Preemption." *Philosophy and Phenomenological Research* 102 (3): 515–30.
- Bond, Rod, and Peter B Smith. 1996. "Culture and Conformity: A Meta-Analysis of Studies Using Asch's Line Judgment Task." *Psychological Bulletin* 119 (1): 111–37.
- Buchanan, Allen. 2004. "Political Liberalism and Social Epistemology." *Philosophy & Public Affairs* 32 (2): 95–130.
- Clifford, W.K. 1877. "The Ethics of Belief." *Contemporary Review* 29: 289–309.
- Feldman, Richard. 2006. "Epistemological Puzzles about Disagreement." In *Epistemology Futures*, 216–36. Oxford: Oxford University Press.
- Gampa, Anup, Sean Wojcik, Matt Motyl, Brian Nosek, and Peter Ditto. 2019. "(Ideo)Logical Reasoning: Ideology Impairs Sound Reasoning." *Social Psychological and Personality Science* 10 (8): 1075–83.
- Goldman, Alvin. 1999. *Knowledge in a Social World*. Oxford: Oxford University Press.
- Hannon, Michael. 2022. "Are Knowledgeable Voters Better Voters?" *Politics, Philosophy & Economics* 21 (1): 29–54.
- Hardwig, John. 1991. "The Role of Trust in Knowledge." *The Journal of Philosophy* 88 (12): 693–708.
- Joshi, Hrishikesh. 2021. *Why It's OK to Speak Your Mind*. New York: Routledge.
- Kahan, Dan, and Ellen Peters. 2017. "Rumors of the 'Nonreplication' of the 'Motivated Numeracy Effect' Are Greatly Exaggerated." *Yale Law & Economics Research Paper No. 584*. <https://doi.org/10.2139/ssrn.3026941>.
- Kahan, Dan, Ellen Peters, Erica Cantrell Dawson, and Paul Slovic. 2017. "Motivated Numeracy and Enlightened Self-Government." *Behavioural Public Policy* 1 (1): 54–86.
- Larsen, Knud S. 1990. "The Asch Conformity Experiment: Replication and Transhistorical Comparisons." *Journal of Social Behavior & Personality* 5 (4): 163–68.
- Longino, Helen. 1990. *Science as Social Knowledge*. Princeton, NJ: Princeton University Press.
- Mercier, Hugo. 2020. *Not Born Yesterday: The Science of Who We Trust and What We Believe*. Princeton, NJ: Princeton University Press.
- Mercier, Hugo, and Dan Sperber. 2017. *The Enigma of Reason*. Cambridge, MA: Harvard University Press.
- Mill, John Stuart. 2008. *On Liberty and Other Essays*. Edited by John Gray. Oxford, New York: Oxford University Press.
- Nguyen, C. Thi. 2018. "Echo Chambers and Epistemic Bubbles." *Episteme*. <https://doi.org/10.1017/epi.2018.32>.

- Nietzsche, Friedrich. 1954. "Human, All-Too-Human." In *The Portable Nietzsche*, translated by Walter Kaufmann. New York: Penguin Books.
- Noelle-Neumann, Elisabeth. 1974. "The Spiral of Silence: A Theory of Public Opinion." *Journal of Communication* 24 (2): 43–51. <https://doi.org/10.1111/j.1460-2466.1974.tb00367.x>.
- Noelle-Neumann, Elisabeth. 1984. *The Spiral of Silence*. Chicago, IL: University of Chicago Press.
- Pauli, Benjamin J. 2020. "The Flint Water Crisis." *Wiley Interdisciplinary Reviews: Water* 7 (3): e1420. <https://doi.org/10.1002/wat2.1420>.
- Persson, Emil, David Andersson, Lina Koppel, Daniel Västfjäll, and Gustav Tinghög. 2021. "A Preregistered Replication of Motivated Numeracy." *Cognition* 214. <https://doi.org/10.1016/j.cognition.2021.104768>.
- Simpson, Robert Mark. 2020. "The Relation between Academic Freedom and Free Speech." *Ethics* 130 (3): 287–319.
- Sloman, Steven A., and Philip Fernbach. 2017. *The Knowledge Illusion: Why We Never Think Alone*. New York: Riverhead Books.
- Sperber, Dan, Fabrice Clement, Christophe Heintz, Oliver Mascaró, Hugo Mercier, Gloria Origgì, and Dierdre Wilson. 2010. "Epistemic Vigilance." *Mind & Language* 25 (4): 359–93.
- Stasser, Garold, Susanne Abele, and Sandra Vaughan Parsons. 2012. "Information Flow and Influence in Collective Choice." *Group Processes & Intergroup Relations* 15 (5): 619–35.
- Stasser, Garold, and William Titus. 1985. "Pooling of Unshared Information in Group Decision Making: Biased Information Sampling During Discussion." *Journal of Personality and Social Psychology* 48 (6): 1467–78.
- . 2003. "Hidden Profiles: A Brief History." *Psychological Inquiry* 14 (3/4): 304–13.
- Street, Sharon. 2006. "A Darwinian Dilemma for Realist Theories of Value." *Philosophical Studies* 127 (1): 109–66.
- Su, Siyan. 2022. "Updating Politicized Beliefs: How Motivated Reasoning Contributes to Polarization." *Journal of Behavioral and Experimental Economics* 96.
- Sunstein, Cass. 2019. *Conformity*. New York: New York University Press.
- Tetlock, Philip E. 2015. *Superforecasting: The Art and Science of Prediction*. New York: Crown.
- Trouche, Emmanuel, Petter Johansson, Lars Hall, and Hugo Mercier. 2016. "The Selective Laziness of Reasoning." *Cognitive Science* 40 (8): 2122–36.
- Vavova, Katia. 2018. "Irrelevant Influences." *Philosophy and Phenomenological Research* XCVI (1). <https://doi.org/10.1111/phpr.12297>.