The Puzzle of Motivated Reasoning

Enoch Yim

"And why do you look at the splinter which is in your brother's eye, but the beam in your eye you do not consider?" (Matthew 7:3)

0.   Introduction

Scholars of various fields have documented that individuals or groups often diverge in their inference in response to the same evidence or information. What explains this phenomenon is the attribution of motivated reasoning.[1] The idea is that, in some cases, the polarization of opinions and beliefs happens because the desire for particular conclusions interferes with the epistemic process of each individual. The psychology of motivated reasoning is in part explicated in terms of identity-protective dispositions in respect to one's political or cultural affiliations, and various interventive methods have accordingly been suggested as to prevent or minimize the negative impacts of this phenomenon.[2] The attribution of motivated reasoning likewise plays an important role in explaining and managing the otherwise baffling phenomenon observed in various sectors in the society. Yet, I suspect that the attribution as a theoretical practice gives rise to a distinctive puzzle *if it is accompanied with two additional assumptions:*

---

[1] Adam Carter & Robin McKenna, "Skepticism Motivated: On the Skeptical Import of Motivated Reasoning," *Canadian Journal of Philosophy*, 50:6, 702-718 (2020); Jon Ellis, "Motivated reasoning and the ethics of belief," *Philosophical Compass*, Vol. 17, Issue 6 (2022); Dan M. Kahan, "Ideology, motivated reasoning, and cognitive reflection," *Judgment and Decision Making*, Vol. 8, No. 4, 407-424, (2013); Thomas Kelly, "Disagreement, Dogmatism, and Belief Polarization," *The Journal of Philosophy*, Vol. 105, No. 10, 611-633, (2008); Charles G. Lord et al., "Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence," *Journal of Personality and Social Psychology*, Vol. 37, No. 11, 2098-2109, (1979).

[2] Kahan, 408; Carolina Flores, "Belief Change & Social Change," (unpublished); C. Thi Nguyen, "Echo Chambers and Epistemic Bubbles," *Episteme* (2018).

unconscious bias and sociality of empirical knowledge. In this essay, I argue that, in order to

solve the puzzle, one must reject either that some biases are introspectively inaccessible or that

the empirical knowledge of facts is social—otherwise the attribution is incoherent.

In Section 1, I define relevant terminologies based on literature works. In Section 2, I

present the puzzle of motivated reasoning. In Section 3 and 4 respectively, I explain the first and

the second part of the puzzle in detail. Especially in Section 4, I will give a reason why it is

difficult to reject sociality of empirical knowledge. In Section 5, I conclude the essay by briefly

commenting on what I think are the upshots and implications of my argument.

1.   Terminologies

In this section, I define relevant terminologies based on literature works. First, I follow

Katia Vavova in defining *irrelevant influences* as "factors that don't bear on the truth of what we

believe."[3] Insofar as I understand it, this definition can stay neutral as to which theory of truth is

to be endorsed. What matters is, given that there is some robust notion of truth, some factors

such as one's cultural or religious background seem to contribute nothing to the truth of the

belief that is acquired. Second, I employ Miriam Schoenfield's definition of *permissivism* that it

is a view that, in some cases, there is more than one rational[4] way to respond to the given

evidence.[5] If permissivism is true, as Schoenfield contends, there are permissive cases in which

---

[3] Katia Vavova, "Irrelevant Influences," *Philosophy and Phenomenological Research*, Vol. XCVI, No. 1, 134-152, (2018), 134.

[4] By *rational* (or *reasonable*), I understand epistemic rationality rather than pragmatic rationality although I also think that one cannot be pragmatically rational without also being epistemically rational or, better put, that—if understood well—epistemic and pragmatic rationality collapse into one without either of them being reduced to the other. However, since this discussion is beyond the scope of this essay, I limit the analysis strictly to epistemic rationality as conventionally understood in contemporary academics.

[5] Miriam Schoenfield, "Permission to Believe: Why Permissivism Is True and What It Tells Us About Irrelevant Influences on Belief," *Nous*, 48:2, 193-218 (2014).

the involvement of irrelevant factors in one's acquisition of a belief does not invalidate the belief because the inference was rational according to the epistemic standard provided by the factors.[6] Naturally, some cases are non-permissive: for some evidence, there is only one rational way to respond to it, or some responses are irrational. Third, by *epistemic standard*, I understand the cognitive disposition guiding or presupposed in the process of making judgments: evidence-collecting practices, inferential patterns, etc.[7] As such, epistemic standards include, but are not limited to, conceptual and linguistic framework although they exclude orectic or affective states or systems because the standard view is that they are not epistemic.

Fourth, Jon Ellis suggests characterizing *motivated reasoning* as "reasoning guided by some goal or end extrinsic to *truth, consistency, and other epistemic goals*."[8] I endorse this definition because it is practically the broadest characterization of motivated reasoning. As such, motivated reasoning is one form of non-permissive irrelevant influences. It is an irrelevant factor because its guidance (involvement in the epistemic process) does not bear on (or necessitate) the truth of the acquired belief. Given that one of the epistemic goals one should aim for is to be rational in judgment based on the criterion provided by his epistemic standard, the involvement

---

[6] Schoenfield, 199.

[7] One may complain that this characterization conflicts with the definition of irrelevant influence because, if some irrelevant factors such as one's upbringing provide epistemic standards and if these epistemic standards in turn provide the criterion for rationality, then those irrelevant factors do bear on the truth of the acquired belief. On the one hand, I think that this points to the need to sharpen the notion of truth and rationality as well as their relationship. Does being rational in itself bear on the truth of the belief? This inquiry is beyond the scope of this essay. On the other hand, even if rationality is a prerequisite for identifying the truth of a belief, one could still distinguish between irrelevant factors and irrelevant factors *proper*. Some irrelevant factors such as culture and religion may not bear on the truth of the belief in the sense that the fact that one has such a background does not *necessitate* that the acquired belief is true (whereas some factors such as evidence seem to directly bear on the truth). Yet, without these factors, the subject would not have developed or been equipped with any conceptual framework to deal with information to begin with. In contrast, some irrelevant factors not only bear nothing on the truth of the belief, but also fail to provide any epistemic standard. The candidates for such irrelevant factors *proper* include, but may not be limited to, desires and hallucinations (although I must add that I think we could still distinguish between desire *proper* and *epistemic* desire which yearns for truth).

[8] Ellis, "Motivated reasoning and the ethics of belief," 7.

of motivated reasoning is non-permissive because motivated reasoning is guided by factors such as desires that encourage irrational responses to the evidence. That is, the response based on motivated reasoning is not counted among the rational responses that are possible. Thomas Kelly argues that motivated reasoning is not necessarily irrational (or unreasonable) because, he thinks, the motivated subject is responding to the total evidence available to *him*. The total evidence gathered by the end may be skewed due to the initial bias. Nevertheless, the conclusion drawn from the evidence is in some sense the best one could do in response to it.[9] However, Kelly fails to distinguish between initial and present responses to available evidence as well as between evidence availability and evidential response. The subject's response to the (total) evidence available *now* may be rational. Yet, that does not mean that the initial response to the first set of evidence (that resulted in the present set) was rational as well. Also, the subject could be responding to the present set either rationally or irrationally. Just because the present set is the total evidence available does by no means guarantee that the response to it is rational. If Ellis' characterization of motivated reasoning is right, in respect to the other terminologies above, the motivated subject's response to the present or the initial evidence should be seen as irrational, i.e., fails to be counted among rational responses that are possible.

Fifth, I define *introspection* as any step in the epistemic activity that does not involve empirical data or induction. As such, introspection may involve logico-linguistic analysis, meditation, or self-reflection. Conversely, extrospective information is any information that could only be reached at by non-introspective methods such as observation and experiment. Lastly, by *unconscious*, I mean that which is undiscoverable or ungraspable via introspection. This definition is intuitive in that, if something is introspectively accessible, then one can be

---

[9] Kelly, 629.

aware of it since he can *think about it*. In contrast, if one cannot introspectively access some

information, then he cannot be aware of its content because it would be absurd to suppose that

there is a sense in which some content could be put in a format accessible by the mind without

recourse to any extrospective method and yet cannot be introspectively captured. Conversely, if

something can be introspectively captured, then its content is accessible by consciousness. (One

can of course *access* unconscious contents via extrospective methods. But the language he uses

to express the contents would be different from were he to access it via introspection.)

2.   The Puzzle of Motivated Reasoning

The puzzle of motivated reasoning is as follows. One cannot accept all three following

statements, i.e., these statements are inconsistent with one another:

(1) The attribution of motivated reasoning to others is a theoretically coherent practice;
(2) Some biases are unconscious (introspectively inaccessible);
(3) The empirical knowledge of facts is social, or the acquisition of empirical knowledge
    relies on (or is mediated by) others.

One may hold up to two of the three statements above, but not all of them at once. In the rest of

this section, I will articulate the general idea of this puzzle; the next two sections will explain (2)

and (3) in detail.

To illustrate the point of my argument, let me start by devising two cases in which one

may attribute motivated reasoning to others.

> *Impartial*. I observe that two groups, Republicans and Democrats, make different
> conclusions in response to the same information provided in regard to the issue of climate
> change. Republicans conclude that the climate change is not driven by human activities
> whereas Democrats conclude the very opposite (by crediting evidence in favor of their
> view or discrediting counter-evidence).[10]

---

[10] This example case is inspired by and based on the study in Dan M. Kahan's "Ideology, motivated reasoning, and
cognitive reflection" (2013) wherein subjects of different political affiliations are divided into different groups of

Partial. On the issue of climate change, I disagree with my peer. I argue that the climate change is driven by human activities whereas my peer argues otherwise. Both my peer and I have access to and are equally familiar with the same information.

These two cases represent the generic cases in which we observe polarizations of opinions and beliefs in response to the same information. In both cases, given that I do not outright dismiss either of the sides as unintelligent, there is the need to explain why these polarizations happen. (Since the cases above are in which two parties have access to the same information *and* neither is attributed with unintelligence, the subject of the inquiry in this essay is limited to what Nathan Ballantyne would call "unresponsiveness-to-compelling-evidence" cases.[11]) One could make at least two hypotheses, one of which is the attribution of motivated reasoning:

> **Epistemic Standards**. The two sides (Republicans and Democrats, my peer and myself) have different epistemic standards due to different backgrounds.

> **Motivated Reasoning**. Either of the two sides is engaging in motivated reasoning.

On the surface, it seems that both hypotheses can equally explain the phenomenon just well. The question is why choose one over the other—more specifically, why choosing the latter results in the puzzle of motivated reasoning presented as above.

2-1.   Partial

Let us see how **Motivated Reasoning** creates the puzzle in *Partial*. In response to my peer's disagreement, I would initially attribute motivated reasoning to my peer rather than to myself; my thought is that my peer's desire for a particular conclusion in favor of his view must

---

various measures and their responses to the same information are observed and analyzed. My case is an overly simplified version. Nevertheless, this version should not be problematic in serving as a good example.
[11] Nathan Ballantyne, "Debunking Biased Thinkers (Including Ourselves)," *Journal of the American Philosophical Association*, Vol. 1, Issue 01, 141-162 (2015), 146-7.

have interfered with his epistemic process. This is perhaps the most natural context in which we feel the need to attribute motivated reasoning to anyone in the first place—as Ellis remarks:

> [The attribution of m]otivated reasoning occurs when we are in disagreement with others. It's brought in as an explanation of the differing conclusion. Where there's agreement, we rarely need or think to explain it. But in contexts of disagreement, we're more likely to be epistemically different, to not appreciate larger, background beliefs. In some cases of disagreement, possible explanations of their reasoning are not easily apparent, though the possibility of motivated reasoning is salient.[12]

So far so good. But a problem arises in respect to certain empirical researches that some biases are unconscious.[13] Specifically, in "The Introspection Illusion," Emily Pronin notes,

> The current theorizing speaks to circumstances in which actors have rich access to introspective information, such as when introspection provides actors with convincing evidence of their good intentions. Importantly, though, it also speaks to circumstances in which actors lack introspective access, such as when introspection fails to reveal to actors the influence of bias on their own judgments because that bias operates automatically and does not leave conscious traces.[14]

> The influence of unconscious bias, whether activated by implicit priming … or by implicit prejudice …, is not accessible via introspection.[15]

The reason why I attribute motivated reasoning to my peer rather than to myself is because I assume that I am right. I have this assumption because, insofar as my own response to the same information is concerned, I am not aware of any bias in myself that might have interfered with my epistemic process. Accordingly, if I am not the biased one, it must be my peer. However, if some biases are unconscious so that introspection is unreliable in tracking them, I could be the one who is engaging in motivated reasoning. If I accept this skepticism, I should entertain the possibility of attributing motivated reasoning to myself.[16] Suppose that I do attribute motivated

---

[12] Jon Ellis, "Motivated reasoning and micro-epistemology," (unpublished), 24.

[13] Ballantyne, 150-6.

[14] Emily Pronin, "The Introspection Illusion," *Advances in Experimental Social Psychology*, Vol. 41, 1-67, (2009), 4.

[15] Pronin, 26.

[16] For the sake of giving a smooth narrative, I presented the skepticism incurred by motivated reasoning as an epistemic one—that *once* I am aware of the skepticism, I get to (or should) entertain the possibility of attributing motivated reasoning to myself. However, the issue is really a metaphysical one. Regardless of whether I am aware

reasoning to myself instead. This means that I do admit that my belief is wrong and my peer is right. If so, however, is there anything for me to disagree with my peer to begin with? After all, I came to disagree with my peer because I assumed that I am right. But, if there turns out to be no strong ground for the possibility that I am right, the disagreement itself becomes rather obsolete. I made **Motivated Reasoning** to make sense of the disagreement. Yet, if this disagreement is nullified, there has been no need to make this hypothesis all along. In turn, if there is no such need, nothing warrants attributing motivated reasoning to my peer in the first place.

To avoid this incoherency (i.e., to attribute motivated reasoning to either my peer or myself), I must be able to do at least one of the two things *if I still accept that some biases are unconscious*: I can either access the fact that I have no bias or assure as a matter of fact that my belief is right without using introspection (i.e., by using empirical methods). I will talk more about this matter in Section 4, but the general point is that this empirical access to the information (about my own biases or the fact of the matter regarding the issue at hand) cannot rely on others; if the empirical knowledge of facts is social, then I fail to make coherent the practice of attributing motivated reasoning. That is, there must be a way someone (expert) is or I am able to acquire knowledge concerning one's own biases or facts directly from evidence. For, if I have to rely on others to tell me whether I am right, I cannot have disagreements with them (as it turns out that either the others actually agree with me or I accept whatever judgment the others make of me). Conversely, if one accepts that empirical knowledge is social, he should

---

of the skepticism, the fact that this skepticism is possible puts me in an unwarranted position for attributing motivated reasoning to others (or myself). This is why this essay focuses on *theoretical* coherency of the practice of attributing motivated reasoning. Even if the possibility that I might be the one engaging in motivated reasoning simply passes by in my thought so that I end up *only* entertaining that possibility, *the fact of a matter that I could be the motivated reasoner* in itself suffices for undermining the ground for the attribution of motivated reasoning.

reject that some biases are unconscious because, given that I must be able to know that I have no bias on my own, I must be able to access this fact via introspection.

2-2.   Impartial

The way in which the puzzle of motivated reasoning figures in *Impartial* is different. In observing two groups polarizing in their opinions and beliefs, I may choose either of the two hypothesis, **Epistemic Standards** and **Motivated Reasoning**, as (it seems) both hypotheses are equally capable of explaining the phenomenon. What drives one to choose the latter would be the view that some cases of polarization such as above (climate change) are *non-permissive*, i.e., there is only one way to rationally respond to the same information. It is the case that the given evidence supports the conclusion either that the climate change is driven by human activities or that it is not, but not both. Thus, at least one of the groups must be wrong, and the reason why this group is wrong is because they are engaging in motivated reasoning.[17] Perhaps, the concept of unconscious bias is employed as to explain *why* the wrong group gets it wrong: if there is nothing wrong with the individual's epistemic process in itself (as the person is simply following the logic of his inference), then what renders it irrational must be some factors, goals or ends that are extrinsic to epistemic goals such as truth and accuracy—the story goes. Unlike in *Partial*, applying unconscious bias to myself (impartial observer) does not undermine the practice of attributing motivated reasoning to either of the group. If two groups contradict each other and if only one of them is right, then the other must be wrong—this is a tautology the truth of which is

---

[17] Here, it is possible that both groups may be engaging in motivated reasoning as even the group that got it right might have been driven by their desires to draw that conclusion. The point is that, given that it is a non-permissive case, the group that got it wrong *must have been* engaging in motivated reasoning.

introspectively accessible. What is *not* introspective accessible, therefore gives rise to the puzzle, is whether it is the case that only one of them can be right.

The problem is at the question of what warrants me to determine that *Impartial* is a non-permissive case. Why is it *not* the case that both of the Republican and the Democrat responses are rational? To bring out the problem more starkly, let us assume that the given information is sufficiently conclusive, i.e., it is not the case that the polarization is simply the result of inconclusive studies.[18] In order to determine that *Impartial* is non-permissive, one must also presuppose that the evidence supports either of the two responses, but not both. This requires that the inferential relationship between the information and the (only correct) response is known or knowable (either by the impartial observer or some expert to whom he can refer).

Can this (empirical) knowledge (of the inferential relationship between the information and the response) be social and yet the impartial observer still maintain that the case is non-permissive? There are two ways in which the answer to this question is negative. First, if we accept that empirical knowledge is social, then it may not cause any trouble for *Impartial*, but in elsewhere, *Partial*. **Motivated Reasoning** must be able to explain both *Impartial* and *Partial*, and if the same explanation does not work in one case (*Partial*), this hypothesis is an *ad hoc*. If it is supposed that the hypothesis is a better fit for only one of the two cases, this asymmetry must be explained. However, the difference between *Impartial* and *Partial* does not seem to be drastic enough to warrant this supposition, for both are dealing with the qualitatively same phenomenon, viz., polarization in response to the same information. Second, accepting that empirical knowledge is social does in fact cause trouble for *Impartial*. Let us say that empirical knowledge

---

[18] If you believe that the researches on climate change are inconclusive, you could substitute it with any other polarizing topic the study of which is relatively conclusive, e.g., vaccination.

is social so that the acquisition of the knowledge of e.g., climate change relies on one another. If so, why is one of the two groups, Republicans and Democrats, (whichever is wrong about climate change) excluded from this epistemic community? Why does not one of them have a say on the issue? That is, who determines that one and not the other of them is wrong or irrational? Perhaps, one must be qualified to be counted as a member of the, viz., scientific community. But who settles the criterion for qualifications? It seems that either there must be a super-community that can give this criterion or the scientific community should be able to settle the issue on its own. In the former case, the question of why one of the two groups is excluded from this super-group arises again, which results in an infinite regress. In the latter case, it is requisite that some insiders (experts) can acquire the empirical knowledge without relying on outsiders (others). Once again, the empirical knowledge of facts must at some level be non-social.

To summarize the argument, one cannot take the attribution of motivated reasoning as a theoretically coherent practice without rejecting either that some biases are unconscious or that the empirical knowledge of facts is social for the following reasons. Accepting the former requires rejecting the latter: some community or I need an empirical way to know that I have no bias, my belief is right, or certain cases are non-permissive without relying on others—as relying on others either (i) undermines the disagreement between me and others which is the prerequisite for attributing motivated reasoning in the first place or (ii) results in an infinite regress in explaining why some others are qualified to determine the permissibility of cases. Accepting the latter requires rejecting the former because, given that I cannot rely on others without undermining the disagreement, there must be a non-empirical (i.e., introspective) way for me to access the fact that I am not engaging in motivated reasoning.

3.   Unconscious Bias

In this section, I will explain in detail why accepting the sociality of empirical knowledge requires one to reject unconscious bias. For the sake of the argument, assume that the acquisition of empirical knowledge relies on others. I have mentioned in the previous section that this condition implies that I cannot genuinely disagree with others including my peer on whose belief is wrong or who has biases. I will consider three possible objections to this part of the puzzle.

3-1.   Relying on others, just not my peer

One may object that the puzzle does not arise because that I have to rely on others does not mean I also have to rely on my peer. My response to this objection is that relying on others *other than* my peer will result in the same problem as the one for *Impartial* which was explained in the previous section. Suppose that I rely on some others, viz., experts (e.g., psychologists, cognitive scientists, etc.) for discovering whether I have biases or not. The question is once again why my peer is not included in this epistemic community, or why I am not willing to grant the same epistemic authority to my peer. If the reason is because these *experts*, but not my peer, are deemed to be qualified for determining the issue, there must be a criterion for qualifications (such as completing some degree programs, professional trainings, etc.). This criterion should be settled either by a super-community (*other than* the expert community) or the expert community itself. Just as in the case of *Informal*, if the criterion is to be settled by the super-community, it is an infinite regress to justify why my peer is not included in that community. If the expert community is self-sufficient, then the empirical knowledge of the presence of biases is at some level not social, i.e., does not rely on *others*. So, *I* may rely on others, and these others may not

even have to include my peer. Yet, some others (experts) cannot rely on any other in yielding the knowledge I need to maintain the disagreement with my peer.

3-2.   The genetic fallacy

It may be complained that I am committing a genetic fallacy by conflating the initial stage in which I attribute motivated reasoning to my peer (disagreement) and the final stage in which I realize that I could be engaging in motivated reasoning (skepticism). It is true that, for my argument, the order in which I come to grasp the force of skepticism is important. First, I must be willing to employ the concept of motivated reasoning and attribute it to my peer. Second, in respect to the other two assumptions (unconscious bias and non-sociality of empirical knowledge), I become skeptical about the attribution. The objection would be that becoming aware of this skepticism does not undermine the initial attribution in disagreement.

The puzzle would be a genetic fallacy only if the ground for skepticism is separate from the possibility of disagreement. However, this is not the case, and accepting the skepticism does undermine the possibility of disagreement wherein this possibility is the very ground on which the skepticism stands. Notice that the skepticism is *not* about the concept of motivated reasoning, but rather about the practice of *attributing* motivated reasoning. Even if I never conceive the need to employ the concept of motivated reasoning, it may as well be the case *as a matter of fact* that my peer is or I am really engaging in motivated reasoning. This is a metaphysical condition whereas the ground for attributing motivated reasoning is epistemic. My attribution of motivated reasoning (to myself, my peer, or both) is valid on the premise that there really is something to disagree about. But, if there in fact is none to disagree about, the attribution itself becomes an empty practice. More strongly put, the word 'motivated reasoning' loses its theoretical meaning.

I grant that, regardless of its theoretical ground, the practice itself may show certain empirical results in reducing polarization (perhaps by promoting intellectual humility[19]). However, then, this would be one exceptional case in which there is a mismatch between theory and practice (such as that classical physics does make rockets work although it does not completely reflect the physical reality of nature at the level of quantum mechanics).

3-3.   Strong vs. weak unconscious

My puzzle presupposes a strong version of unconscious as it is defined in Section 1: what it is for x to be unconscious is that it is introspectively inaccessible. It may be argued that this is too strong of a requirement and that some unconscious biases may be introspectively accessible if enough efforts are given. That is, introspection may in general be unreliable in tracking biases, but it does not follow that biases are *in principle* inaccessible via this method. Perhaps, this is so. However, that would be just to redefine what it means for x to be unconscious. I do grant that there is no puzzle if biases are introspectively accessible (even if using introspection to track them requires much more effort than one can practically input so that it may be more efficient to use empirical methods). But, if we endorse a weaker definition for unconscious, I think that the word 'unconscious' loses its explanatory power.

In the previous section, I briefly mentioned that one of the reasons why the concept of unconscious bias is employed in the first place may be to explain why some individuals or groups diverge in their inferential conclusions, given that they are not outright unintelligent.[20] If the diverging individuals as intelligent agents are interested in epistemic goals, what makes them

---

[19] Carter & McKenna.

[20] In regard to this explanatory function of the concept of unconscious bias, I have in mind what Ballantyne calls 'abstract theories of bias'; Ballantyne, 152.

behave against their (epistemic) interests? The hypothesis here is that these individuals are not aware of their own biases. Others may help such an individual become aware of his biases by presenting various empirical evidence, and this effort may render the bias *conscious*. However, that conversion of self-knowledge then depends on non-introspective methods such as others presenting empirical evidence, etc. If the individual did not need any empirical intervention to begin with, he would have just discovered the biases via introspection in the first place. Yet, if this is possible, then the explanation for why this individual did diverge is unsatisfactory. Thus, for such an explanation to be successful, biases must be conceived as *strongly unconscious*. Here, one may counter-argue that the motivated subject may be able to discover his own biases himself without relying on interventions by others. If this is correct, then it means there is a non-introspective way for an individual or group to discover inner biases that does not rely on others, in which case the assumption that empirical knowledge is social must be rejected.

4.   Sociality of Empirical Knowledge

Granted that one cannot accept that empirical knowledge is social without rejecting that some biases are unconscious in the strong sense, the natural question is why one should accept the sociality of empirical knowledge. There is of course the linguistic version of the thesis that the acquisition of empirical knowledge relies on others, which is advocated by Hillary Putnam, viz., the hypothesis of the division of linguistic labor.[21] However, just because some words are meaningful only in reference to what others know due to specialization does not necessitate that all words are meaningful in that manner. Analogously, that we have to rely on others for some

---

[21] Hillary Putnam, "The Meaning of 'Meaning'," *Language, mind, and knowledge*, Vol. 7, 131-193, (University of Minnesota Press, 1975), 144-6.

empirical knowledge does not mean that we always have to rely on them. If this is the case, it may be argued, there is no reason to accept that (every kind of) empirical knowledge is social, in which case the puzzle of motivated reasoning is not an interesting one.

I do grant that the thesis that empirical knowledge is social is not conclusive, nor is there a general consensus on this issue. In fact, some pieces of empirical information even seem to be immediately graspable and require no reliance on others. For instance, I *know* that I have two hands, that I am sitting on a chair in front of a desk, that I do not (or my peer does) behave in the manner that involves biases, etc. Based on Vavova's work on irrelevant influences, one may posit that, if there is non-social empirical evidence that one is or is not biased, he is justified in attributing motivated reasoning to himself or the other. This posit is based on Vavova's Good Independent Reason Principle (GIRP): "To the extent that you have good independent reason to think that you are mistaken with respect to *p*, you must revise your confidence in *p* accordingly—insofar as you can."[22] Motivated reasoning is a non-permissive type of irrelevant influence, and the acquisition of some empirical (self-)knowledge does not rely on others—if so, then one has "good independent reason" for attributing motivated reasoning to oneself or the others. The problem with this counter-argument is that it is not so clear what this "good independent reason" consists in. In addition, if we are serious about the sociality of empirical knowledge, there is a sense in which even the seemingly immediately graspable information such as that I have two hands is mediated. Rejecting this view requires a strong non-social theory of knowledge that is unrivaled or some knockout argument against the social theory. In the rest of this section, I will articulate why the notion of "good independent reason" is ambiguous and (briefly) present an argument in favor of social theories of knowledge.

---

[22] Vavova, 145.

4-1.	Good independent reason

	Take Vavova's own example that Cohen observes that the Oxford graduates tend to
accept the analytic/synthetic distinction whereas the Harvard graduates tend to reject it.[23] Does
this give Cohen a good independent reason to revise his confidence or attribute motivated
reasoning? As Schoenfield suggests, the polarization could be explained away by **Epistemic
Standards**.[24] But what if Cohen discovers further, as Vavova puts it, that the administrations at
both universities flipped the coin to decide whether to feed their students with the food the side-
effects of which include holding false philosophical beliefs? Or what about another example by
Vavova that a test subject just heard from his experimenter that the experimenter flipped a coin
to decide whether to stimulate the subject's brain to have false experiences?[25] If we assume that
these cases are indeed non-permissive, then there is a good independent reason for Cohen or the
test subject to doubt his beliefs. But why should they correlate the drugged food or the brain
stimulation with coming to hold beliefs that they would have otherwise not come to hold? What
is the good reason or evidence for this correlation? If we assume that there is no correlation, then
Epistemic Standards seem to be able to explain away even these cases. One could still argue
that the mere possibility that those things could be correlated with false beliefs provides some
reason to be skeptical. However, Vavova distinguishes between mere possibility and probability
of error.[26] If the mere possibility of error should concern us, Cartesian skepticism or the brain-in-
a-vat scenario should as well, which requires the principle that Vavova thinks we should reject,

---

[23] Ibid.
[24] Schoenfield.
[25] Vavova, 143, 146.
[26] Ibid., 143-4, 148-50.

viz., No Independent Reason Principle (NIRP): "To the extent that you *fail* to have good independent reason to think that you are *not mistaken* with respect to matters like *p*, you must revise your confidence in *p* accordingly—insofar as you can."[27] Thus, one either accepts NIRP (in addition to GIRP) or suppose that there is an empirical way to make the correlations (insofar as he accepts unconscious bias). If NIRP should be rejected (as Vavova suggests), then it should be determined whether the empirical method is social or non-social. Likewise, what counts as "independent good reason" hinges on the question of the sociality of empirical knowledge.

4-2.   Social theories of knowledge

Non-social theories of knowledge would hold that there is a way for an individual to acquire knowledge or belief directly from evidence without relying on others. What I have in mind as an example for this view is some form of Russellian sense datum theory although there may be other variations. As much as the debates on these theories are not conclusive, I think that there is a way to make a good case for social theories of knowledge. In the rest of this sub-section, I will sketch out what I believe to be one viable version of such theories.

There is a sense in which Robert Brandom's Hegel and Saul Kripke's Wittgenstein overlap in that both interpretation works identify normativity as what gives meaning or content to our thoughts and languages and that both locate the source of normativity in sociality.[28] It would be beyond the scope of this essay to go into details of these works. Nevertheless, some outlines of their overlapping account would suffice for grounding my argument. Both *Brandegel*

---

[27] Ibid., 148.
[28] Robert B. Brandom, *A Spirit of Trust: A Reading of Hegel's* Phenomenology, (The Belknap Press of Harvard University Press, 2019); Saul A. Kripke, *Wittgenstein: On Rules and Private Language*, (Blackwell Publishing, 1982).

and *Kripkenstein* identify as brute impulses (desire[29] and inclination[30] respectively) the

psychological factor in one's going by one pattern of behaviors rather than another. So, at the

bottom, it is merely due to the brute impulse that one responds to the evidence on climate change

or the "+" sign in certain ways (by, e.g., uttering, "I do not believe it!" or writing, '2 + 2 = 5').

What gives content, meaning, or conceptuality to such an impulsive pattern of behaviors so that

there are correct and incorrect (linguistic or doxastic) responses is recognition (*Brandegel*)[31] by

or agreement (*Kripkenstein*)[32] with others. That is, it is only when our utterances or behaviors are

constrained by one another are they subject to standards of correctness and incorrectness; in turn,

it is only in respect to such normativity do their patterns have conceptual or inferential contents.

     Kripkenstein remarks that there is no agreement outside a community.[33] Similarly,

Brandegel contends that recognition is held at the level of community.[34] Thus, what gives

conceptual or inferential content to and imbue one's thoughts and actions with correctness and

incorrectness is their sociality. That is, my knowledge (justified true belief)—that climate change

is driven by human activities *as well as* that I have two hands—relies on others not merely in the

shallow sense that I have to refer to others (experts) for confirmations, but also in the deeper

sense that its content and acquisition (i.e., entrance into the *space of reasons*) are mediated by

others. The reason why normativity requires constraints by others is because one can only be

properly *corrected* by others. As David Landy puts it, the idea behind this view is that there are

---

[29] Brandom, 235-61.
[30] Kripke, 87-91.
[31] Brandom, 262-312.
[32] Kripke, 91-8
[33] Ibid., 96.
[34] Brandom, 262-465.

only three candidates for the source of normativity: objects, self, and others.[35] The criticism on

immediacy and the Given in the philosophical tradition from Hegel to Wilfrid Sellars[36] shows

that (sensorial intakes of) objects themselves cannot constrain our thoughts. If one is to make

rules for himself as he goes by, the rule-following is whimsical, which is absurd. Therefore,

normativity must be coming from constraints by others (or one another). If this view is correct,

then conditions that make up knowledge (conceptuality, inferentiality, justifiability, etc.) are

social, i.e., rely on and are mediated by the social interactions with others.

5.   Conclusion

        The main upshot of my argument is that, if we want to employ **Motivated Reasoning**

(i.e., consider as coherent the practice of attributing motivated reasoning), then we must do one

of the two things: either reject unconscious bias or sociality of (empirical) knowledge, to do

which one must provide a knockout argument against the strong concept of unconscious bias or

social theories of knowledge. Particularly, rejecting the social theories of knowledge leads one to

various other discussions in epistemology and its related fields where the debates on the nature

of knowledge and belief (as well as normativity and necessity) are, as far as I see it, still ongoing

inconclusively. Without settling these epistemological issues, the status of the legitimacy of

Motivated Reasoning is at best indeterminate. (Likewise, I think that the real problem is the

rivalry between the social and the non-social view of knowledge, or differently put, the tension

between the foundationalist intuition that some knowledge is direct or unmediated and the

---

[35] David Landy, "Hegel's Account of Rule Following," *Inquiry: An Interdisciplinary Journal of Philosophy*, 51:2, 170-193, (2008).

[36] Wilfrid Sellars, *Empiricism and the Philosophy of Mind*, originally published in 1956, (Harvard University Press, 1997).

challenge posed by the philosophical criticism on immediacy and the Given.) Overall, insofar as there are good reasons (as, I believe, provided in this essay) to doubt the explanatory power of a weaker concept of unconscious bias or the idea of immediate knowledge or the Given, one must accept the challenge put forward as the puzzle of motivated reasoning.

One may attempt to avoid the puzzle by maintaining unconscious bias and accepting sociality of knowledge, but abandoning the practice of attributing motivated reasoning. In that way, one may promote stopping the attribution of motivated reasoning altogether. There are at least two problems with this escape route. First, as I mentioned in Introduction, the concept and the attribution of motivated reasoning play various important roles in explaining and managing polarizations happening in various sectors. Thus, giving up its practice and concept would be costly, if not impossible. Second, discarding **Motivated Reasoning** leaves us with **Epistemic Standards** (if these two hypotheses exhaust the options), and this hypothesis is not without any problem. Consider the following questions. Given that there are polarizations, how do we decide which ones are permissive or non-permissive? Even if the case is permissive, could we determine that one epistemic standard is in some sense *better* than the others? If so, what would be the criterion if it cannot be rationality (since both standards are *rational*)? Or, is there some meta-standard by which we can examine the rationality of each epistemic standard? If not, what should we make of the disagreements on various permissive topics? Are epistemic standards incommensurable? Does incommensurability imply cultural relativism? Or, is there some (non-comparative) criterion by which we can judge between epistemic standards? These questions are some of which must be answered if we give up attributing motivated reasoning.

Let us assume, for the sake of supposition, that we have good reasons for maintaining unconscious bias and accepting sociality of knowledge. Does it give us the conclusive motivation

for giving up the attribution practice? Yes and no. I say, yes, because—as it is demonstrated in the essay—the puzzle of motivated reasoning and the three theses (theoretical coherency of attributing motivated reasoning, unconscious bias, and sociality of knowledge) are inconsistent *insofar as* the concept of motivated reasoning is defined in this essay, i.e., as a form of non-permissive irrelevant factor. More explicitly put, as long as we employ *this* concept of motivated reasoning (which seems to capture what many scholars in relevant fields have in mind), the puzzle will remain. But I also say, no. The social theories of knowledge (if true) suggest that the nature of knowledge and belief presupposed in our foundationalist intuition is rather problematic. Accordingly, once the true nature of knowledge and belief is understood, perhaps our concept of motivated reasoning may change. In turn, this version of motivated reasoning might coherently fit together with the other two theses (or hypotheses) and still does its part in explaining and managing polarizations just as well. Whether this (interdisciplinary?) project is viable is an open question, but I think that it is something worth investigating into.

References

Ballantyne, Nathan. "Debunking Biased Thinkers (Including Ourselves)." *Journal of the American Philosophical Association*, Vol. 1, Issue 01. 2015.

Brandom, Robert B. *A Spirit of Trust: A Reading of Hegel's Phenomenology*. The Belknap Press of Harvard University Press: 2019.

Carter, Adam & Robin McKenna. "Skepticism Motivated: On the Skeptical Import of Motivated Reasoning." *Canadian Journal of Philosophy*, 50:6. 2020.

Ellis, Jon. "Motivated reasoning and micro-epistemology." Unpublished.

———. "Motivated reasoning and the ethics of belief." *Philosophical Compass*, Vol. 17, Issue 6. 2022.

Flores, Carolina. "Belief Change & Social Change." Unpublished.

Kahan, Dan M. "Ideology, motivated reasoning, and cognitive reflection." *Judgments and Decision Making*, Vol. 8, No. 4. 2013.

Kelly, Thomas. "Disagreement, Dogmatism, and Belief Polarization." *The Journal of Philosophy*, Vol. 105, No. 10. 2008.

Kripke, Saul A. *Wittgenstein: On Rules and Private Language*. Blackwell Publishing: 1982.

Landy, David. "Hegel's Account of Rule Following." *Inquiry: An Interdisciplinary Journal of Philosophy*, 51:2. 2008.

Lord, Charles G. et al. "Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence." *Journal of Personality and Social Psychology*, Vol. 37, No. 11. 1979.

Nguyen, C. Thi. "Echo Chambers and Epistemic Bubbles." *Episteme*. 2018.

Pronin, Emily. "The Introspection Illusion." *Advances in Experimental Social Psychology*, Vol. 41. 2009.

Putnam, Hillary. "The Meaning of 'Meaning'." *Language, mind, and knowledge*, Vol. 7. University of Minnesota Press: 1975.

Schoenfield, Miriam. "Permission to Believe: Why Permissivism Is True and What It Tells Us About Irrelevant Influences on Belief." *Nous*, 48:2. 2014.

Sellars, Wilfrid. *Empiricism and the Philosophy of Mind*, originally published in 1956. Harvard University Press: 1997.

Vavova, Katia. "Irrelevant Influences." *Philosophy and Phenomenological Research*, Vol. XCVI, No. 1. 2018.