

PHI 315: Philosophy of Mind

2020

Prompt:

What is the problem of introspection? How might we respond?

You might focus on just one aspect of the problem and one aspect of the response.

What it is like to Believe

Nimra Nadeem

Introspection is the distinctive means, whatever they may be, by which we come to know certain facts about our own minds. **Distinctive** because we have means of coming to know facts about our own minds which are distinct from the means we have of coming to know facts about the external world including the minds of others. (Cox, Lecture Notes) The following **epistemological principle makes** introspective knowledge problematic: “It is wrong always, everywhere and for anyone to believe anything on insufficient evidence.” (Clifford) Keeping this in mind, the problem of introspection is formulated as follows:

P1: There are introspectively knowable mental facts about our own minds

P2: If there are introspectively knowable facts about our own minds, then all such facts are knowable only on the basis of evidence.

P3: If there are introspectively knowable facts about our own minds, then not all such facts are knowable only on the basis of evidence.

Clearly, all the propositions cannot be true at the same time. (Cox, Lecture Notes) My challenge in this paper will be directed at proposition 3.

For clarity, I define the following terms.

Foundational Evidence: knowledge that serves as its own evidence. For e.g. if we accept perception as a form of foundational evidence then when I perceive a red apple, the perception is itself evidence that I believe that I perceive a red apple.

Regress Stopper: X is a regress stopper if it qualifies as foundational evidence, and thus serves as a solution to the problem of justificatory regress.¹

Phenomenal quality: A state X has a phenomenal quality if there is something it is *like* to be in X.

There are other definitions relevant to my argument, which I will introduce along the way.

¹ The term regress stopper is in a way synonymous to the term foundational evidence. However, I introduce and subsequently use the term regress stopper to emphasize that our purpose behind identifying certain kinds of evidence as foundational is to avoid the justificatory regress problem.

I begin by presenting the problem of justificatory regress and its foundationalist solution. The purpose of this section is to establish the premise upon which my later argument stands: that being phenomenal is a necessary and sufficient condition to serve as foundational evidence. Next, I introduce Crispin Wright's formulation of the hierarchy of beliefs we hold about ourselves. I then modify Wright's approach to present my full argument, summarized as follows:

We have a rich array of beliefs about ourselves. All such beliefs fall into two categories: (1) immediate beliefs and (2) attitudinal beliefs.² The first kind serve as foundational evidence due to their phenomenological character, and thus we know these kinds of beliefs merely by virtue of having them. The second kind is known through inference from the first kind.

The epistemological principle mentioned above gives rise to the problem of justificatory regress. A belief B is justified only if it's supported by some evidence X. In order to have evidence X, you must be justified in believing X. In order to be justified in believing X you must support it with some evidence Y. In order to have evidence Y, you must be justified in believing Y... and so forth. There's a common foundationalist solution to this regress problem which proposes that evidence based on perception can serve as *foundational evidence*, what I refer to as a *regress stopper* in my later argument. [cite] The thought is that perceiving puts us in contact with the world in a way that doesn't require independent justification of the perception.

If we are persuaded by this foundationalist approach, we concede that a fundamental regress stopper is needed and that perception serves as such a regress stopper. An important premise for the argument I present in this paper is as follows:

P4: What serves as the necessary and sufficient condition for perception to qualify as a regress stopper is its *distinctive phenomenal quality*.

By distinctive phenomenal quality, I mean that there is something it is *like* to perceive something, which is uniquely associated with the experience of perception. This proposition is supported by Michael Huemer's *phenomenal conservatism*, where he reasons as follows:

If undefeated appearances are not a source of justified belief, then how is one to avoid skepticism about the external world, the past, values, abstract objects, and so on?

² I define these terms later.

Unless this challenge can be met, we would be wise to place our trust in the appearances.... (Tucker, 349)

Here it is in particular the *phenomenal quality* of perception that serves as the justification for its qualification as foundational evidence. I believe this reasoning can be intuitively understood in two ways. Recognizing that we are committed to skepticism,³ if we allow skepticism about that which is phenomenally experienced ('apparent' in Huemer's terms), then it is hard to imagine how anything could ever be known or believed, for lack of sufficient evidence. Another way of thinking about it is this: we want to not be skeptics. The simplest way to do so is by using phenomenal knowledge as foundational. Simple explanations are better than convoluted ones, so we accept this hypothesis.⁴

Now we turn to the problem of introspection and what could ground it. I begin by presenting Wright's views. In his piece *Self-Knowledge: The Wittgensteinian Legacy*, Crispin Wright identifies the source of the problem of introspection to be the phenomenon of **avowals**, which he defines as *the phenomenon of authoritative, non-inferential self-ascription*. (page 14) He claims that there are 2 different kinds of avowals that any response to the problem of introspection needs to account for. **Phenomenal avowals**, which can be intuitively understood as avowals describing some phenomenal experience of the individual. They comprise of examples like "'I have a headache', 'My feet are sore', 'I'm tired'..." (page 14) These kinds of avowals satisfy three properties. They are **groundless**, which, in simple terms, means that in response to such an avowal, asking "'How can you tell?'"—is always inappropriate." (p. 14) They are **strongly authoritative**, i.e. a doubt about such a claim has to be a doubt about the sincerity or the understanding of the one making it. (14) They exhibit **transparency**, i.e. there is something "absurd about a profession of the form, 'I don't know whether P'" (15) **Attitudinal avowals** can be broadly understood as avowals describing one's *attitude* towards something or someone. They consist of examples like "'I believe that term ends on the 27th', 'I hope that noise stops soon'..." (15) These avowals also exhibit groundlessness and transparency, although

³ This is the point of foundationalist theories, to fight skepticism, so I think this is a fair assumption to make

⁴ I will not delve further into this discussion due to the length constraints of this paper. But whether or not the reader is convinced of this argument so far, the fact remains that experts in the field of foundationalist epistemology (notably Huemer) have argued so, and as such functioning under this premise still allows my argument to remain academically relevant.

they only exhibit what he calls **weak authority**, i.e. “there is space for relevant forms of self-deception or confusion.” (17) Wright distinguishes between attitudinal avowals and cases of self-interpretation. Attitudinal avowals are more foundational in that they describe “non-inferential knowledge of a basic range of attitudes and intentionally characterized responses”, (16) whereas **self-interpretation** is characterized by inferential knowledge *based on* the knowledge from phenomenal and attitudinal avowals. Since self interpretation is inferential, responses to the problem of introspection need not identify its justificatory source, though they must explain the source of justification for phenomenal and attitudinal avowals.

I will now present my view and in the process show where I disagree with Wright. I claim that there are two kinds of beliefs we may hold. **Immediate** beliefs, knowledge of which is groundless, transparent and strongly authoritative as per Wright’s terminology. The second kind of beliefs I call **attitudinal beliefs** in that they represent our attitudes towards something, similar to Wright’s description of the content of attitudinal avowals. I propose that *all attitudinal beliefs* rely on inference and interpretation of the more basic, immediate kind of beliefs. My crucial point of departure from Wright’s view is this: I claim that there is no real category of attitudinal avowals. Self-ascriptions either represent immediate beliefs, in which case they *are* strongly authoritative (and knowledge of these is justified due to their phenomenal character which I will elaborate on next) and are indeed *avowals*, i.e. non inferential and authoritative. Or self-ascriptions represent interpretive beliefs, in which case they are inferences derived from more basic self-knowledge and thus do not require an independent source of justification.

Instead of using Wright’s terminology of **avowals**, I think it may be more helpful to think in terms of mental facts. There are two kinds of mental facts that we can know about ourselves. The first are **phenomenal mental facts**. These are the facts that we come to know in a *groundless, transparent and strongly authoritative way*. This may sound very similar to Wright’s phenomenal avowals, but in my view this is a much richer category of mental facts. They include not only regular sensory and perceptual experience, but also the *phenomenal experience of immediate beliefs*. The second kind of mental facts are simply those that represent our interpretive beliefs, so I will refer to these as **interpretive mental facts**.

Phenomenal mental facts include facts about perception, but go beyond just regular sensory perception. They include “raw feelings” of all sorts, for example, “I feel sad”, “I feel happy”, “I feel nauseous,” which are not directly associated with regular sensory perception but are nevertheless uncontroversial in terms of their phenomenal quality. But this category of mental facts also includes immediate beliefs, those that are immediately transparent to the mind by virtue of being held. This is where my category of phenomenal mental facts is wider and richer than the subject matter of Wright’s phenomenal avowals. Let’s say it’s around 9pm. Someone asks me, “do you believe that it is 9pm?”. I would have a definitive answer to the question the moment the question is asked, which is transparent in the sense that I can’t not know whether I believe it is 9pm. It is also strongly authoritative in the sense that asking “are you sure that is what you believe?” doesn’t make sense.⁵ And less obviously, perhaps, my answer is also groundless. One could say that my belief that I believe it is 9pm is grounded in my perception that the clock ticked 9. But it is plausible that I hold the belief that it is not 9pm even if the clock ticked 9. Consider the case of lucid dreaming. I could be perceiving all sorts of things, and yet not believe that they reflect reality. Thus my answer is groundless such that asking me “how do you know” is not a valid question. I may be able to answer how I think I reached that belief (I believe that it is not 9pm because I am aware I am dreaming) but I can’t quite answer how I know that this is what I believe. A final crucial characteristic of immediate beliefs is that they have a *distinct phenomenal character*, that is there is something it is *like* to hold an immediate belief which is distinct from what it is like to perceive evidence for that belief.

Consider the example of John Nash from the movie A Beautiful Mind. He develops paranoid schizophrenia which results in extremely deceptive auditory and visual hallucinations. By the end of the movie, Nash learns to reason his way into distinguishing between hallucinations and reality. For example, he realizes that his best friend’s niece, Marcee, is a hallucination because she never gets old. The fact is, despite perceiving the presence of Marcee, Nash believes that she is not really there. There is certainly something it is *like* to be John Nash which is distinct from what it was like to be him when he used to perceive Marcee in front of him and believe her to be real. If someone were to ask the older Nash whether he believes

⁵ Note that yes someone may respond by saying “are you sure?” but this would make sense only when my belief itself that it’s 9pm is being challenged, not the belief that I believe.

Marcee is real, Nash would have an immediate, transparent, groundless and strongly authoritative answer to the question. Note that Nash might have *reached* the belief that Marcee is a hallucination based on inferential reasoning, but the belief itself, once formed, has a distinct phenomenal quality which is what makes it an immediate belief.

I now ask the reader to return to the proposition I made at the very beginning, that a distinctive phenomenal quality of X serves as the necessary and sufficient condition for X to qualify as a regress stopper, i.e. as foundational evidence. Since I have shown that all phenomenal mental facts, including immediate beliefs have a distinct phenomenal quality, I can now claim that all phenomenal mental facts serve as regress stoppers. That is, phenomenal mental facts, like perception, can serve as the foundational basis for our justificatory hierarchy. In particular, for the case of immediate beliefs, that you believe is all that it takes to justify that you believe the belief.

My argument can be formally written as follows:

P5: X can be considered a regress stopper if and only if X has a distinctive phenomenal quality.

P6: If X can be considered a regress stopper, X requires no independent source of its justification.

P7: Immediate beliefs (and phenomenal mental facts in general) have a distinctive phenomenal quality.

C1: Immediate beliefs (and phenomenal mental facts in general) can be considered regress stoppers.

C2: Immediate beliefs (and phenomenal mental facts in general) require no independent source of their justification.

The last part of my case is to show that knowledge of the second kind of mental facts, what I referred to as interpretive mental facts, is not epistemologically controversial. This is where my view is opposed to Wright's. Interpretive mental facts include only what I defined as attitudinal beliefs.⁶ I claim that *all attitudinal beliefs* are inferential, based on the more basic self

⁶ These two terms, attitudinal beliefs and interpretive mental facts refer to the same thing. The use of two different things may seem a bit superfluous but just to establish a clear distinction

knowledge of phenomenal mental facts. I will challenge Wright's own example of Jane Austen's *Emma* to show why this is the case.

Wright quotes the scene where Emma reflects on her own emotional reaction to Harriet professing love for Mr. Knightley. Upon this reflection, Emma comes to the realization that she is in love with Mr. Knightley herself. Wright points out that when Emma interprets her reaction to Harriet's declaration "there is an avowable ground —something like 'I am disconcerted by her love for that man and, more so, by the thought that it might be returned—which is a datum for, rather than a product of, self-interpretation.'" (page 16) This avowable ground is what he calls attitudinal avowals and he claims that these are distinct from phenomenal avowals because they are only 'weakly authoritative' in that self deception is a possibility and thus they serve as a different kind of "basic self-knowledge" which is non inferential. I disagree that the *basic datum* Emma uses in her process of self interpretation is her "disconcert for Harriet's expression of love." Rather, I think Emma uses her self-knowledge of her phenomenal mental states to first arrive at the belief that she is disconcerted by Harriet's love for Mr Knightley. She subsequently uses this attitudinal belief to further infer that she is in love with Mr Knightley.

I think that whenever there is room for self-deception/confusion, it is a case of self-interpretation based on phenomenal mental facts. In the case of Emma, her self interpretation is based on the phenomenal mental facts she comes to know when she thinks of Harriet's love for Mr Knightley - for example, she feels her heart rate rise, she feels the raw feelings of sadness, distress, and anger. Does she believe that she feels bad? Yes, on the basis of an immediate belief that she feels bad. If she is sad, she cannot possibly deny that she feels sad. *Why* she feels sad is a matter of interpretation. So her belief that she is disconcerted by Harriet's love is the case of inferential reasoning.

In this way, cases of self-knowledge about beliefs which are not immediate are based on inference from either the knowledge of immediate beliefs or the knowledge of some other phenomenal mental fact. Thus Wright's category of attitudinal avowals is not a distinct category, and doesn't require a separate source of justification.

between the two kinds of mental facts and the two kinds of beliefs, I thought it was worth using an extra term.

My full response to the problem of introspection is thus as follows. Introspective knowledge is based on phenomenal mental facts, which serve as foundational evidence due to their phenomenal character as per P1. Introspective knowledge of all other mental facts is inferential, based on these phenomenal mental facts, and thus is epistemologically uncontroversial.

References

- Clifford, William K. "The ethics of belief." *First Published* (1877).
- Cox, Ryan. "Lecture notes - PHI 315: Philosophy of Mind". Princeton University. (2020).
- Tucker, Chris, ed. *Seemings and justification: New essays on dogmatism and phenomenal conservatism*. Oxford University Press, 2013.
- Wright, Crispin. "Self-knowledge: the Wittgensteinian legacy." *Knowing our own minds* (1998): 13-46.