Why Think that Belief is Evidence-Responsive?*

Carolina Flores

1   Introduction

The orthodox view in philosophy is that belief is constitutively evidence-responsive. I will offer a novel argument for the following version of the orthodox view:

> **Evidence-Responsiveness Constraint**: Necessarily, if S believes that p, then S has the capacity to respond to evidence bearing on the belief that p by rationally updating their belief that p.[1]

My starting point is the Sellarsian idea that the concept of belief functions to mark the space of reasons. The basic idea is the following. The mind is a complex jumble of mental states that form various patterns. Users of mental state concepts parse these patterns into attitude types in ways responsive to our interests. One interest we have is picking out attitudes that we can change by offering evidence, and that are governed by epistemic standards. I contend that the concept of belief is partly for picking out such attitudes. The result is that there is a close connection between an attitude being a belief and being evidence-responsive.

In this paper, I will flesh out and defend this picture. In §2, I will describe the methodology that I will employ to argue for the Evidence-Responsiveness Constraint. §3 to §6 will go over the steps in this methodology. §3 articulates the idea that a central function of the concept of belief is to mark the space of reasons. §4 moves from this to the Constraint. §5 shows that the resulting view includes both evidence-resistant and animal beliefs. §6 considers additional benefits of the view: it illuminates the normative practices surrounding belief, helps us distinguish belief from secondary cognitive attitudes such as imaginings, and brings a novel perspective to the psychology of belief revision and to ethical and political issues involving evidence-resistance.

1. This view assumes, as standard, that belief is an attitude toward a proposition. But it can be adapted to allow belief to have non-propositional contents by extending a notion of relevant evidence to non-propositional contents (with accompanying standards for rationally responding to such evidence).

2   Methodology

What do we want the concept of belief for? What purpose does classifying things along belief-lines serve? Given the answers to these questions, which things in the world does "belief" pick out?

Asking these questions follows in the footsteps of *function-first epistemology*, a recent trend that examines the function of our epistemic concepts to determine the nature of epistemic kinds. I will focus on Craig (1991) and Hannon (2018)'s version of function-first epistemology. Here is how Edward Craig describes the method:

> We take some *prima facie* plausible hypothesis about what the concept does for us, what its role in our life might be, and then ask what a concept having that role would be like, what conditions would govern its application...and then see to what extent it matches our everyday practice with the concept as actually found (Craig 1991, 2–3)

This can be broken down into three stages (Hannon 2018). First, we generate a hypothesis about what the target concept is for. For instance, Craig and Hannon propose that the concept of knowledge serves to mark good informants. Why? Because identifying good informants is a central need for creatures who need true beliefs and who must rely on testimony for them. It would make sense for us to use a central concept like "knowledge" here.

Once we have a starting hypothesis about the function of a concept, we can derive its application conditions. For instance, when it comes to knowledge, Craig and Hannon take the perspective of an agent—and of a community of agents—trying to identify good informants.

We them test the proposed application conditions to see if they sufficiently match our intuitive classifications. A perfect match with all intuitions is not required. As long as we have sufficient match to vindicate the hypothesized function for that concept, we can reject intuitions that do not fit the account to preserve the claim that the concept has the function identified. This can be seen as a form of reflective equilibrium, where we balance out a claim about the function of the concept and intuitions about cases.

One might worry about this function-first methodology being overly focused on the *concept* of belief, as opposed to on *beliefs* themselves. Why should we think that our concepts are a guide to reality? After all, our concepts often carve reality poorly, due to lack of knowledge: think here of the concept of "gold" before we knew enough about molecular structure to distinguish gold from pyrite.

I agree that our concepts can be a poor guide to reality, and that we need to proceed with caution moving from concepts to reality. We need to consider whether the concepts we have are worth keeping. In some cases, like in the case of gold, we need to revise our concepts and carve the world in new ways in light of our knowledge.

For this reason, I add an extra step to the methodology that Craig and Hannon propose: considering whether we should continue to classify things in this way. In particular, does our scientific knowledge of the mind imply that we should revise our taxonomy, much as we revised how we classify metals in light of advances in

chemistry? And similarly, because our scientific knowledge is not all that matters to classification, is this way of classifying attitudes helpful for normative theorizing, and from an ethical and political perspective?

Despite the need for this additional step, I don't think that we can eliminate thinking about the functions of the concept of belief and just think about the science of belief to determine the nature of beliefs (as Quilty-Dunn and Mandelbaum (2018) suggest). This is because the concept of belief is not just at the service of scientifically explaining behavior: we also think and talk about beliefs in contexts of evaluation and social regulation. Following Schroeter and Schroeter (2015), addressing "What is x?" questions requires balancing out different interests we have in employing the concept of x. Simply looking to science risks failing to do justice to these other interests.[2] What I want to do in this paper is precisely to highlight some of these other interests— ones connected to real-life epistemic assessment and regulation—and consider their consequences.

Let us take stock. This is the method I will apply to thinking about belief:

1. What are the central function(s) of the concept of belief? Answer: one central function is to mark the boundaries of the space of reasons (§3).

2. What application conditions follow? Answer: for participants to be in the space of reasons with respect to a belief, that belief must be evidence-responsive (§4).

3. Do these application conditions sufficiently match our intuitions, enough to vindicate the initial hypothesis about the function of the concept? Answer: Yes; in particular, the account includes animal beliefs and evidence-resistant beliefs (§5).

4. Do we have reason to continue classifying attitudes as beliefs in this way, in light of scientific, theoretical, and ethical considerations? Answer: Yes on all counts (§6).

3   The concept of belief marks the space of reasons

Arguably, there are different concepts that we refer to with the ordinary language term "belief." Most notably, ordinary English speakers primarily use "belief" to refer to religious faith and the like, not to ordinary factual beliefs, such as the philosopher-favorite belief that there is beer in the fridge (Heiphetz et al. 2021). In contrast, like virtually every philosophical account of belief, my account counts run-of-the-mill factual beliefs as paradigmatic beliefs. I am trying to get a better understanding of belief

---

2. The concept of belief is often held to have additional functions. For instance, some hold that the concept of belief functions to mark the end of inquiry (Friedman 2019), or commitment (Basu 2021, Buchak 2014) or stability (Leitgeb 2014) in one's take on the world. Does the target concept of belief really serve these functions? If the concept of belief serves these functions, are they derived from one another or independent, each generating constraints on the application conditions of the concept? These are interesting questions that I can't discuss here.

in *that* sense (what Westra (forthcoming) calls "epistemic belief"), not in the ordinary language sense.[3]

My starting hypothesis is that the concept of belief functions to mark the space of reasons. This draws inspiration from Sellars's famous point:

> In characterizing an episode or a state as that of *knowing*, we are placing it in the logical space of reasons, of justifying and being able to justify what one says." (Sellars 1956, 169)

This idea is at the root of a prominent 20th century tradition of thinking that there is a deep connection between the normative (on the one hand) and mental content in general, or belief in particular (on the other). Interpretationists (Davidson 2004, Dennett 1981) take the Sellarsian point to indicate a connection between the normative and mental content: for attitudes to have content, they must be part of a sufficiently rational web. Normativists take the normativity to which Sellars points to be located in the attitude *type* of belief, not in content. Belief is in the state of reasons in that it is by its essence a state to which epistemic norms apply—even if they don't satisfy those norms (Brandom 1994, Boghossian 2003, Helton 2020, Velleman 2000, Wedgwood 2002; see McHugh and Whiting 2014 for an overview).

I find this Sellarsian idea intriguing, and, given its distinguished history, I am not the only one. I think there is something worth keeping about it. That said, I depart from the intellectualist ways in which this idea has been developed. In particular, I don't want to tie the space of reasons to the ability to justify what one says, which I think is too cognitively demanding for belief.[4]

Indeed, I think there is a more pared-down way of thinking about the space of reasons that is quite intuitive. Engaging in the space of reasons is a matter of offering evidence (epistemic reasons) with respect to a claim under discussion, while taking (however implicitly) epistemic norms to be applicable standards for how the attitude should change in light of that evidence. The space of reasons thus has a normative dimension: the standards of rationality apply. This normative dimension is causally implemented in how we interact with one another, and these standards are made manifest in our practices of criticism and accountability.

The relevant contrast class to engaging in the space of reasons includes sheer pressure or bullying, affective coloring, appeals to emotions, and reframing in ways that favor one's view. In these cases, one does not offer direct evidential support for a hypothesis. Instead, one offers the agent practical reasons to find a way to believe that hypothesis (in the case of bullying or pressure), or causes the agent to preferentially attend to or search for such evidence.

The central, Sellarsian-inspired idea I start from, then, is this: the concept of belief functions to mark participants in the space of reasons. More precisely, it functions to mark that the subject with that attitude toward *p* is eligible to engage in the space of

---

3. In this sense, belief encompasses both full belief and degrees of belief or credence. The account I develop leaves open how to distinguish these two attitude types.

4. Following Sellars, I am here focusing only on *epistemic* reasons. One could conceive of the space of reasons more expansively, to include reasons for action.

reasons with respect to *p*.[5]

Whether someone is in the space of reasons with respect to a particular claim is highly significant. When you and your interlocutor are in the space of reasons with respect to *p*, you can enter the social game of giving and responding to reasons about *p*. This opens the door to rational engagement and joint deliberation. These are important social practices in their own right: they matter for community-building, joint action, and affiliating with one another. They are also epistemically important practices. Limited social agents like us need to reason together to arrive at an accurate view of reality. The importance of these practices makes this a distinction that agents like us have an interest in demarcating.

Further, if belief marks the space of reasons in this sense, we have a simple explanation for why we epistemically assess *beliefs*. According to my view, the concept of belief functions to single out attitudes which are in a space where epistemic standards are the operative ones. And, if the subject is in the space of reasons with respect to an attitude in this sense, then there is a practical point to epistemic standards: these standards guide the evidence we offer, and attitudes can in principle be brought to meet those standards.

Even with all this on the table, one might object that marking the space of reasons is not a core function of the concept of belief, not one which should guide our theorizing. Why not start with the common-sense hypothesis that the concept of belief functions to explain and predict behavior?[6]

This is a function of the concept of belief. However, it does not suffice to individuate belief. Other attitudes also function to predict and explain behavior (what of desires?). So my opponent's claim about the function of the concept of belief needs refining. They need to specify which role in explanation and prediction the concept of belief functions to mark. Once they do so, they will also owe an account of why we would have a concept that picks out *that specific role in explanation.* The proposal that "belief" marks the space of reasons allows us to see, out of all the attitudes that play a role in the explanation and prediction of behavior, which ones "belief" picks out, and why we want to pick those out.

Less concessively, I worry that this objection parochially centers our interests in explanation and prediction to the detriment of assessment and regulation, and a detached, third-personal stance over the involved stance of a participant in interpersonal interaction. Such a concern with explanation appears supreme to those of us who think in the terms of modern science. But it might be less universal than it seems, and downplay the importance of interests that are central to interacting and not just passively observing each other—such as the interests we have in rational engagement.[7]

---

5. This parallels the claim that the concept of knowledge functions to mark good informants—specifically, that someone is eligible to be a good informant with respect to a claim. Note, also, that this is compatible with people being in the space of reasons with respect to attitudes that are not beliefs, such as suspensions.

6. Thanks to Eric Schwitzgebel for pressing this point.

7. Cf. Craig 1991, 5.

4 From the space of reasons to evidence-responsiveness

The next step is to derive this hypothesis' application conditions.[8] When would we mark someone as eligible for participation in the space of reasons with respect to a claim p?

Suppose that you are trying to decide whether to engage with someone in the space of reasons. This requires that they are available to you here and now and that there are channels of communication open between you, otherwise you can't convey evidence to them. Further, they need to be open to responding to you offering evidence in ways that you take to be rational within the time frame and energy you have for interacting with them, and you need to be able to tell that this is the case.

We might summarize these considerations by saying that $S2$ is a participant in the space of reasons for $S1$ with respect to p just in case:

    a $S2$ is available to $S1$ here and now

    b There are open channels of communication between $S1$ and $S2$

    c $S2$ is likely to respond to evidence on p offered by $S1$ by updating $A(p)$ in what $S1$ takes to be rational ways within the (time, etc.) constraints $S1$ has for the interaction

    d $S1$ can detect that condition (c) is met.

These are not the application conditions for the concept of belief. The concept of belief is not responsive to the granularity of each agent's needs. In saying that someone believes that p, we are conveying to others that the person is in the space of reasons with respect to p. These others may vary arbitrarily with respect to their contexts, purposes, views of rationality, and practical constraints. Our own parochial constraints are not relevant to the application conditions of the concept. We need to objectivize, i.e. to reduce the ties of a concept to the needs of particular users (Dancy 1992).

We can immediately scratch conditions (a) and (b) above. All that we need to tell others that someone is in the space of reasons with respect to a claim p is: if we were to have an open channel of communication, engaging in the space of reasons about p would be available. Similarly, the analysis above included a detectability condition (condition (d)), which will not figure in an objectivized version: at most, that other conditions are met must be in-principle detectable.

Condition (c) is the key condition in the account. Here it is again:

> $S2$ is likely to respond to evidence on p offered by $S1$ in what $S1$ takes to be rational ways within the (time, etc.) constraints $S1$ has for the interaction.

This also needs objectivization. First, it is relativized to the persuader's take on what is rational. What matters from a social point of view is not whether the person would change their mind in ways some specific observer likes. Instead, what matters is whether they would do so in ways a suitably idealized community member would

---

8. The exposition in this section is indebted to chapter 2 in Hannon 2018.

endorse—i.e., in rational ways. Hence, the condition needs to be modified to mention "responding to evidence in rational ways."

Second, we want to identify attitudes that are available for targeting with evidence in some general sense, not in response to a particular agent in a particular context.

At first, it might seem that we want to communicate that an attitude is such that *anyone* in *any* context would change it by offering evidence. But such a proposal incorrectly turns "belief" into an achievement term, a mark of being a good and collaborative participant in the space of reasons, not of meeting minimal standards for inclusion. The resulting view is extensionally inadequate. Many ordinary beliefs are not this easy to change by offering evidence. Indeed, we expect engagement in the space of reasons to require active effort, and the prospects of success to be variable.

At the opposite extreme, we might think that any attitude that it is *possible* to change by offering evidence counts. But mere possibility will not do. The fact that someone would adjust their attitude toward p to evidence by accident in one freak scenario does not warrant conveying to others that they are a genuine participant in the space of reasons with respect to p. In ascribing beliefs, we offer a guarantee to others, and mere possibility of change is not enough to offer such a guarantee. Instead, I suggest that we offer a guarantee of adjustment to evidence in suitable contexts in virtue of some feature of the attitude itself.

What we want, I think, is for the agent to have the *capacity* to rationally respond to the evidence offered.[9]

When agents have such a capacity, there is a feature of the attitude that is responsible for producing rational responses to evidence. This makes offering evidence robustly available as a strategy for modulating that attitude. At the same time, capacities are fallible. In particular, they might be masked: other factors (e.g., distraction, desires to hang on to one's beliefs, social pressure, or difficulties suppressing automatic thoughts) might interfere with their functioning (Bird 1998). When capacities are masked, agents will not revise their beliefs in response to counter-evidence. As a result, it might be very difficult or time-consuming to change their mind by offering evidence.[10]

Further, capacities have a function. In this case, the function is rationally responding to evidence. This matters for accounting for the normativity of the space of reasons, that is, for the fact that epistemic standards apply to attitudes for which participants are in the space of reasons. Tying beliefs to capacities to rationally respond to evidence helps us account for this feature of the space of reasons. This is because standards of rational revision are appropriate standards to apply to attitudes regulated by evidence-responsiveness capacities. Those standards are met when the capacity succeeds, i.e. correctly exercises its function. There is a match between the capacity's function and what our normative practices require.

---

9. Drawing on Schellenberg (2018)'s account of perceptual capacities, I detail what such capacities involve in Flores ms..

10. Why not require more for inclusion in the space of reasons, e.g. sufficient likelihood of success by offering evidence? This proposal faces several problems, including the difficulty of specifying such a threshold and of any such threshold ending up separating attitudes with very similar cognitive profiles. Further, it is a key fact about human interaction in the space of reasons that its success is highly context-dependent, and, in particular, depends on the persuader. Conveying information about some threshold of likelihood being met fails to recognize this.

If all the above is right, we can revise condition (c) as follows:

> *S* is in the space of reasons with respect to their attitude *A*(*p*) if and only if *S* has the capacity to respond to evidence bearing on *p* by rationally updating *A*(*p*).[11]

If "belief" marks the space of reasons, then we get the following necessary condition on belief:

> **Evidence-Responsiveness Constraint**: Necessarily, if *S* believes that *p*, then *S* has the capacity to respond to evidence bearing on the belief that *p* by rationally updating their belief that *p*.

5   The Evidence-Responsiveness Constraint is extensionally adequate

To vindicate the idea that "belief" marks the space of reasons, I will now argue that the Evidence-Responsiveness Constraint tracks our practices of belief ascription.

Ordinary factual beliefs clearly satisfy this constraint. Beliefs about whether it is raining outside, or there is beer in the fridge, are unproblematically evidence-responsive. They adjust to evidence in immediate and effortless ways.

Still, many beliefs are evidence-resistant. Can my account allow for them? And can it account for belief in animals, given that animals don't appear to participate in the space of reasons? I will argue that the answer to both is "yes." My proposal largely matches our intuitive and scientific classification practices using the concept of belief.

5.1   Animal beliefs

Connecting belief to the space of reasons might suggest excluding non-human animals, as well as infants and people with severe cognitive disabilities. After all, the paradigm case of interacting in the space of reasons is a reasoned back-and-forth between two adults. Non-human animals do not participate in such interactions.

I think that if a view of belief precludes animal belief, it should be rejected. We must do justice to the ordinary practice of ascribing beliefs to animals, and to the role this plays in our best science of animal cognition (Andrews 2020). Here I depart from others who have emphasized connections between belief and the space of reasons, such as Davidson (1982) and McDowell (1996), who embrace the exclusion of animal belief.

Happily for my view, we have reason to think that many non-human animals have the capacity to rationally adjust beliefs in response to evidence. This is because rationally responding to evidence does not, in my view, require complex cognitive capacities.

Adjusting a belief in light of the evidence in a way that matches norms of rationality does not require meta-cognitive reflection. Indeed, it can be done sub-personally,

---

11. Strictly speaking, the Evidence-Responsiveness Constraint only requires the "only if" direction.

outside conscious awareness. It also does not require linguistic capacities. An agent needs linguistic capacities to possess (and hence respond to) evidence conveyed through verbal testimony. But these capacities are not needed to respond to, e.g., perceptual evidence.

Indeed, it is plausible that even very simple animals have evidence-responsive attitudes. For instance, bees seem to construct mental maps of their environment and update these maps in response to new information: they communicate with one another about the distance and direction of sources of pollen using the bee dance, and they incorporate in their mental maps the information they receive through the dance (Gould et al. 1988). Plausibly, such updating in response to this new information suffices for meeting the Evidence-Responsiveness Constraint.

At this point, one might worry that the constraint lets in any attitude whatsoever. Bu this is not true. Some animal behavior is the result of fixed, innate routines. For instance (cf. Carruthers 2004), caterpillars follow the light to climb trees to find food. This might seem to suggest that they have the capacity to rationally respond to evidence about what direction is up. But this is not the case. When artificial light is provided at the bottom of trees, caterpillars climb down and starve to death. They lack the capacity to adjust their behavior in response to new information. Their behavior is dictated by a simple mechanism that makes it the case that, when more light enters one eye than the other, the legs on that side move slower, causing the animal to turn toward the light. Such behavior is *not* the result of attitudes that meet the Evidence-Responsiveness Constraint.

Similarly, attitudes that are regulated exclusively via associative learning are not evidence-responsive. Associative learning requires cumulative experiences of association. In contrast, evidence-sensitive learning can be a one-shot matter, with a single item of evidence forming or extinguishing the relevant attitude.

In sum, my view liberally ascribes beliefs to animals, while still setting lower bounds on belief. This counts in favor of its extensional adequacy.

5.2 Evidence-resistant beliefs

We are unlikely to revise political, moral, or religious beliefs (Markus 1986, Leeuwen 2014), beliefs in theories we are committed to (Chinn and Brewer 1993), and beliefs about ourselves and our talents (Gilbert 2006). Can my account do justice to these cases?

The answer is 'yes'. The central reason for this is that capacities can be masked, resulting in evidence-resistant belief. My view contrasts with interpretationist views that exclude attitudes that sufficiently resist counter-evidence.

To show that my view can accommodate evidence-resistance, I will focus on the most extreme case of evidence-resistant belief: clinical delusions, which I have argued in Flores 2021 to be evidence-responsive. If my view counts even delusions as beliefs, there should be few obstacles to counting in other real-world evidence-resistant beliefs.[12]

---

12. In my Flores ms., I argue for the inclusion of ordinary evidence-resistant beliefs in more detail, examining the central mechanisms behind evidence-resistance.

There are many positive reasons to think evidence-responsiveness capacities are present in delusions. Here are two points. Patients put substantial effort into avoiding counter-evidence (Freeman et al. 2001). This suggests that, if they were to acquire counter-evidence, it would be hard to resist abandoning the delusion (f that evidence would have no effect, why avoid it?). More decisively, in successful CBT, patients actually respond to counter-evidence to their delusion, and thereby come to abandon it. CBT's relatively high success rate (Lincoln and Peters 2018) indicates that (many) patients have the capacity to rationally respond to such evidence.

Though these capacities are present, they are systematically masked in delusions. For example, many patients have persistent altered perceptual experiences which yield apparent evidence for their delusions (Ellis and Lewis 2001). And patients are often highly motivated to maintain their delusion, either because it is a pleasant one, or because abandoning it would require the painful realization that something has gone seriously amiss (Bentall et al. 2001). Further, many have general difficulties overriding cognitive biases; if they improve at doing so, they have an easier time abandoning their delusions (Moritz et al. 2014).

Delusions, then, result from the layering of capacities to rationally respond to evidence and masks on those capacities. I think it is highly plausible that we can understand other cases of evidence-resistance (e.g., religious faith, ideological stances) along similar lines.

Further, my view counts in evidence-resistant beliefs while accounting for our ambivalence with respect to whether many such cases count as beliefs. There are live debates on whether delusions (Bortolotti 2009), religious faith (Leeuwen 2014), and ideological or political stances (Hannon 2021) count as beliefs. Indeed, some have argued that these are borderline cases of belief: neither determinately beliefs nor determinately some other attitude (Schwitzgebel 2001, Tumulty 2014).

My account predicts such ambivalence. Due to the possibility of masks, it is hard for the naked eye to tell whether these attitudes meet the Evidence-Responsiveness Constraint on belief. Nevertheless, there is a determinate fact of the matter about whether they are beliefs, one which psychology can help us discover. Their borderlineness is epistemic, not metaphysical.

Despite its overall extensional adequacy with respect to real-world cases and fit with intuitions of ambivalence, the account excludes fully rigid or unrevisable attitudes that in other ways behave exactly like beliefs.[13] One might think this is decisive: if an account cannot fit all intuitions, it must go.

I can see the intuitive pull of classifying these as beliefs. However, the function-first approach I have followed gives us principled grounds on which to discount these intuitions. If my argument is correct, agents are not in the space of reasons with respect to these claims. Marking these attitudes as beliefs would misuse the concept

---

13. It also excludes the attitudes that Smithies et al. (2022) talk about: attitudes characterized by a deep feeling of conviction, but which do not in other ways share features of ordinary beliefs. The points below apply to these cases.

An interesting further question is what the account entails about beliefs with credence 0 or 1, which are not rationally revisable. I think there are two ways to go here: either rejecting that they are beliefs, and justifying that based on the reasons in the next paragraph, or rejecting the claim about rational revisability and saying that it is irrational to have these extreme credences.

of belief, falsely implicating that rational engagement is on the table. Further, there are independent reasons to think that resembling beliefs in other ways—in inferential promiscuity, action-guidance, and connections to sincere assertion—is not sufficient for belief. In particular, it seems that lively fantasies can play these roles, yet they are not beliefs (Helton 2020).

We do not need to count these cases as beliefs to correctly classify real-world evidence-resistant beliefs. Hence, satisfying intuitions about these cases is a desideratum that is plausibly outweighed by doing justice to the function of the concept of belief I have identified. We might instead think of such attitudes as cognitive analogues of fixed perceptual priors (Stone et al. 2009), or as non-belief hinge certainties (Coliva 2016).[14]

6   Why we should keep this concept of belief

I have argued that the concept of belief picks out evidence-responsive attitudes. But perhaps, as both Schwitzgebel (2021) and Zimmerman (2018) have argued, we should reform the concept of belief, and do so in such a way that fully evidence-insensitive beliefs count.

Against this, I will highlight benefits of conceiving of belief as marking the space of reasons, and thus as evidence-responsive. To preserve these benefits, we should keep employing this concept of belief.

I have already alluded to some of its advantages: accounting for animal belief and evidence-resistant belief, doing justice to our ambivalence about whether attitudes such as delusions are beliefs, and explaining the significance of ascribing beliefs. I will now showcase additional benefits: for explaining our practices of epistemic assessment and interpersonal epistemic regulation; for contrasting beliefs with other attitudes (e.g. imaginings); for theorizing about the psychology of belief; and for interpersonal interaction and political contexts.

6.1   Explaining normative practices

In §3, I noted that connecting belief and the space of reasons allows us to explain the fact that we apply epistemic standards to beliefs. Now that we have the Evidence-Responsiveness Constraint on the table, we can flesh out a fuller picture of the epistemic normativity of belief, i.e., of why epistemic standards (standards of rational responsiveness to evidence) apply to belief.

On the view developed here, beliefs are governed by evidence-responsiveness capacities: capacities that have as their *function* rationally responding to evidence. Func-

---

14. Alternatively, you could take this function-first methodology to yield a feature that figures in the prototype of belief, not a necessary condition. Such a view would be able to include fully evidence-resistant beliefs while preserving a connection between the category of belief and evidence-responsiveness. I am open to this idea, but more interested in exploring the Evidence-Responsiveness Constraint to (a) show that it can handle real-world evidence-resistant beliefs, (b) preserve the claim that classifying an attitude as a belief commits us to subjects being in the space of reasons, and (c) help us develop a clear-cut taxonomy and empirical models for the study of belief revision, one which allows for a single unified model in which to study all belief revision.

tions set standards of success: the capacity succeeds when it fulfills its function (in this case, when the belief in question rationally responds to evidence), and it fails otherwise. So beliefs are successful to the extent that they are evidence-responsive. Rational evidence-responsiveness is a standard that governs belief.

The concept of belief, then, picks out attitudes that are normative in the weak, functional sense of normativity.[15] It is the normativity of success and failure, not of objective standards that we ought to care about. (Consider: if the function of a murderer is to murder, this does not mean that there is some substantive norm enjoining murder). This means that the view I develop here is friendlier to naturalism than normativist (McHugh and Whiting 2014) views, which hold that normativity (in a thick sense) is essential to belief. In my view, drawing on Millikan 1986, the only normativity that is part of the nature of belief is the normativity of function. And the normativity of function is familiar from biology and psychology.

Though this functional normativity is in the nature of beliefs themselves, this is a picture that leaves room for our social practices in explaining the normativity of belief. The key idea is that our practices of engagement in the space of reasons embody a notion of "working" for belief: in particular, they embody standards for when beliefs appropriately adjust to evidence.[16] It is because we care about such normative standards that we use the concept of belief to select attitudes for which this notion of "working" makes sense, namely, attitudes that are sustained by evidence-responsiveness capacities.

The fact that the attitudes we select with our concept of belief have this functional profile explains how epistemic standards can have causal force. Epistemic standards, on this view, hook into our causal practices of trying to change others' minds. Normative epistemic reasons (evidence) are also causes of belief revision, when they function as input to successfully exercised evidence-responsiveness capacities. This gives epistemic standards a practical point: they guide us in selecting evidence when trying to change others' minds.

This said, our practices of epistemic regulation, in some cases, extend beyond this causal-push-and-assessment picture. Our engagement in the space of reasons with other adults often takes an angstier form, involving the reactive attitudes (Strawson 1962): we get angry, disappointed, or upset when others don't respond rationally. We may voice this to them and attempt to get them to see the force of these norms and to therefore respond to evidence accordingly.

The view I have articulated does not explain the angsty form which our engagement in the space of reasons with other adults takes (though it can perhaps be expanded to encompass it). This is a strength of the view. It reflects the fact that these richer normative practices do not apply to *any* belief in *any* cognitive system.[17] For

---

15. Sullivan-Bissett (2017) defends a similar view of the normativity of belief. We differ in that Sullivan-Bisset thinks that some beliefs are not produced by systems with epistemic functions, and, correspondingly, that they are outside the realm of epistemic normativity.

16. For more on starting from our practices of engagement in theorizing about epistemic norms, see Flores and Woodard forthcoming.

17. Indeed, an attitude can be epistemically evaluable without the subject who has that attitude being in the purview of epistemic oughts (cf. the distinction between evaluative and deontic). Plausibly, for an agent to be in the purview of ought-claims, it must be appropriate to praise or blame them for whether they comply (in the absence of excuses).

instance, they do not apply to non-human animals who have beliefs (§5). As such, explaining their applicability requires us to look beyond belief.

There are many proposals in the literature about the application conditions for such practices. All require more than the Evidence-Responsiveness constraint. For instance, according to McHugh (2013), it is required that the belief be formed or sustained by a mechanism that is receptive and reactive to epistemic reasons. The agent must sufficiently often recognize the reasons she has and react to these by acting as she takes them to recommend (McHugh 2013). In contrast, evidence-responsiveness as I described it does not require agential recognition of reasons. Similarly, McGeer and Pettit (2002) argue that responsibility for beliefs requires flexible self-regulation in response to evidence, as opposed to merely having one's beliefs regulated by evidence.

In sum, connecting belief and evidence-responsiveness explains the kinds of normative practices that beliefs are a part of merely in virtue of being beliefs. At the same time, it helps us see what additional ingredients might be needed to account for normative practices that involve responsibility.

6.2 Taxonomy

A standard point in favor of the view that belief is constitutively evidence-responsive is that it allows us to distinguish beliefs from secondary cognitive attitudes, such as imaginings and non-doxastic acceptances (Velleman 2000, Leeuwen 2014). Here I will show that my version of the view preserves this benefit.

Let's assume that the following plausible principle:[18]

> If a subject S has the capacity to respond to evidence bearing on p by rationally updating their belief that p, then, whenever S receives evidence bearing on p in appropriate conditions and without masks on this capacity, S rationally updates their belief that p.

This yields a test for whether an attitude is a belief. Specifically:

> If there is some possible situation where S receives evidence bearing on p in appropriate conditions and without masks on their evidence-responsiveness capacities at play and does not rationally update A(p), then A is not a belief.[19]

We can use this test to show that my view correctly excludes imaginings and other secondary cognitive attitudes from the belief category.

In my view, then, prescriptive norms on belief do not apply universally, so we should not look to the nature of belief to account for the normativity of belief in the thick sense of accounting for prescriptive norms on believers. Similarly, I disagree with Helton (2020)'s ought-implies-can argument for an evidence-responsiveness constraint on belief. I reject the claim that, necessarily, if a subject believes that p, they ought to revise that belief in rational ways in light of counter-evidence. This ought-claim does not apply to all believers. I endorse the weaker claim that, necessarily, if S believes that p, then S's belief that p is evaluable for whether it is rationally adjusted in light of counter-evidence.

18. I go into more detail on this principle in Flores ms..

19. "Rationally updating" should be read as referring to the norms for belief updating.

Consider imagining that you are on a luxury tropical vacation while grading in your apartment (and receiving decisive counter-evidence from your senses to the claim that you are on vacation). Your relevant beliefs remain anchored to reality: you believe that you are grading in your apartment and not on a luxury tropical vacation. This reflects the fact that you successfully exercise your capacities to rationally respond to evidence. But the successful exercise of these capacities, made manifest in your beliefs, leaves the imagining intact. By the principle above, the imagining does not count as a belief.

This generalizes. It is possible (in fact, common) to fail to revise one's imaginings in accordance with the evidence in good conditions for the exercise of one's evidence-responsiveness capacities.[20] Similarly, non-doxastic acceptances, mere hypotheses, and so on often persist even when beliefs remain tethered to the evidence.

Indeed, *not* being evidence-responsive may be constitutive of imagining and accepting. In diametric opposition to beliefs, these attitudes play the cognitive role of helping us get away from reality and explore alternative ways things could be. Imagining and accepting may be subject to a decoupling mechanism that insulates them from counter-evidence (Leslie 1987, Perner 1991). If that is right, then there are irremovable barriers to evidence-responsiveness capacities operating on these attitudes, in the sense that, if those barriers were removed, the attitude would cease to be an imagining or acceptance.

Another strength of the view developed here is that it suggests a potentially fruitful framework in which to study belief revision: the layered model of belief revision.

According to this model, belief revision is regulated by three layers: evidence-responsiveness capacities, internal masks, and external conditions.[21] This model agrees with traditional epistemology that beliefs are fundamentally regulated by evidence, in that beliefs are constitutively in the province of evidence-responsiveness capacities, and belief change is directly driven by evidence, and only indirectly by non-evidential factors. The second layer consists of non-evidential, mind-internal factors that mask evidence-responsiveness capacities, such as motivational factors, imaginings, and emotions. These shape the evidence that is the input to evidence-responsiveness capacities. The third layer is constituted by mind-external factors: the evidence available in an agent's environment as well as factors such as social affiliations, environmentally-produced cognitive load, and dominant cultural frames, which can shape elements of the second layer, functioning as masks on evidence-responsiveness capacities.

---

20. Two objections. First, it is possible for an agent to adopt a principle where they deliberately modulate their imaginings in accordance with the evidence they receive. In response, it remains the case that there is some possible situation where they fail to do so, viz., one where they are not motivated to modulate their imaginings in this way.
    Second, the possibility of agents adopting such a principle suggests that some imaginings do involve the capacity to rationally respond to evidence, where these capacities have different success conditions than those involved in the capacities required for belief. Specifically, the success conditions include wanting to adjust one's imaginings to the evidence. My response is that these are not the capacities that are relevant to whether an attitude counts as a belief. Those are not contingent on one deciding to respond to evidence. The relevant capacities involve direct regulation of attitudes by evidence.
21. Beliefs might also change in ways that have nothing to do with changes in evidence: think, for example, of forgetting. This model does not apply to such cases.

This model can be employed to describe findings in the psychology of belief maintenance and revision, such as findings about motivated reasoning (Kunda 1990) and the cognitive role of emotions (Lerner et al. 2015). Ideally, it would be developed into a computational model of belief revision. Such a model would make precise the variables that affect how an agent responds to evidence, and allow us to make concrete predictions about responses to evidence in different contexts.

Further, if we understand epistemic rationality in a Bayesian way, the Layered Model allows us to combine the insights of Bayesian theories of belief updating (Tenenbaum et al. 2011) with those of cognitive dissonance theory (Cooper 2007). Bayesians claim that belief revision is always a matter of performing Bayesian updating on one's beliefs. In contrast, according to cognitive dissonance theory, agents sometimes revise their beliefs in ways that contradict Bayesian predictions, *increasing* their degrees of belief in the light of counter-evidence (Mandelbaum 2019).

In my view, the Bayesian story about belief updating captures what happens when the second layer is sufficiently inactive: (bounded) Bayesian updating on the agent's evidence. Findings that Bayesianism struggles to explain (such as the belief polarization effect) are the result of the second layer's operation, that is, of factors such as motivation affecting the exercise of evidence-responsiveness capacities. In this picture, Bayesian machinery is always present and plays an important role. But we need resources beyond Bayesianism to describe belief updating.

The view I develop here is potentially fruitful in suggesting an improved framework for for psychology. This counts in favor of my view. This is significant for a view that takes a Sellarsian starting point. The Sellarsian line is that attitude ascription is "not giving an empirical description of that episode or state" (Sellars 1956, 169), contrasting the empirical with the normative. This skepticism that belief is a psychological state that can be empirically studied pervaded many views derived from Sellars' (e.g. Davidson 1982).

Instead, the picture I have articulated is one which brings together two diverging strands in analytic philosophy: the Quinean idea that mental kinds are natural kinds that figure in our best science of the mind, and the Strawsonian idea that mental concepts (which pick out these kinds) are thoroughly enmeshed in practices of interpersonal regulation. This allows us to do justice both to normative and causal dimensions of belief, instead of myopically focusing on one to the detriment of the other (or while incorrectly rejecting the other dimension).

6.3   Ethics and politics

Finally, connecting belief and the space of reasons engenders a perspective that may help us resist politically and ethically damaging narratives about human nature.

First, it helps us resist narratives on which people are irredeemably irrational. On these narratives, people make bad decisions—both on personal and political levels—because they have bad beliefs. More strongly, they make bad decisions because they are irredeemably bad believers—stupid, childish, irrational, ignorant, and either unable to tell fact from fiction or totally uninterested in the truth.

This narrative has played an important role in attempts at explaining the last decade's global turn toward illiberal, anti-democratic politics. And it has been used

to support attacks on democratic forms of government (Brennan 2016, Achen and Bartels 2016). If ordinary people are irredeemably bad at reasoning, then widespread political participation will not secure social goods, and so elites should make collective decisions in a way that bypasses the people's beliefs.

The account of belief I develop provides resources for resisting this narrative. Irrational belief is not a sign of the inability to rationally respond to evidence, but of masks on rationality. The anti-democratic conclusions above are unwarranted. When faced with irrational belief, it is a live option to identify what masks on rational capacities are at play, and to find ways to remove or circumvent them.[22] This might not be easy to do, and it might require large-scale change: nonetheless, this point encourages us to consider causal levers that respect the values of auto-determination and joint deliberation.

At the other extreme, my view helps us resist the Panglossianism of finding ways to interpret even the most bizarre beliefs as rational. Arguably, we see applications of this general outlook in the work of many Bayesians about belief updating, such as Dorst (2023) and Lieder and Griffiths (2020). It also shows up in Begby 2021, Kelly 2008, and Nguyen 2021. Outside the analytic tradition, this outlook tends to take the form of relativism about rationality.

I agree that we can be too quick to ascribe irrationality, especially in the face of deep disagreement. But following this Panglossian strategy wholesale defangs epistemology, making epistemic norms powerless at collectively leading us to the truth.

In contrast, my view leaves room to acknowledge frequent irrational belief. And it does not require us to abandon stringent standards of rationality in favor of ones that count most people as rational most of the time. As such, it acknowledges the intuitive point that there may be many occasions for legitimate criticism of others' bad epistemic behavior—and for encouraging reform. Here as elsewhere, accepting our fallibility is also making space for aspiration and improvement. By unmasking and employing our rational capacities, we can hope to move towards a shared, accurate view of the world.

7 Conclusion

The Sellarsian idea that "belief" marks the space of reasons is unfashionable. This idea has been developed in ways that support false claims such as that non-human animals do not have beliefs, that beliefs require language, and that beliefs cannot be scientifically studied.

I develop this Sellarsian idea in a new direction, one which preserves the idea that belief is normatively inflected while bringing it into the fold of empirical study. The view I argued for connects belief and capacities to rationally respond to evidence, doing justice to the role of belief in epistemic interaction. At the same time, it includes evidence-resistant belief and animal belief, and it paves the way for better causal models of belief revision, ones which might help us avoid simplistic and noxious takes on human rationality.

---

22. This echoes Thi Nguyen's call for work on *hostile epistemology* Nguyen (2023), the study of environmental features that exploit our cognitive vulnerabilities to lead of epistemically astray.

Works Cited

Christopher H Achen and Larry M Bartels. *Democracy for realists.* Princeton University Press, 2016.

Kristin Andrews. *The animal mind: An introduction to the philosophy of animal cognition.* Routledge, 2020.

Rima Basu. A tale of two doctrines: Moral encroachment and doxastic wronging. In Jennifer Lackey, editor, *Applied Epistemology*, pages 99–118. Oxford University Press, 2021.

Endre Begby. *Prejudice: A study in non-ideal epistemology.* Oxford University Press, 2021.

Richard P Bentall, Rhiannon Corcoran, Robert Howard, Nigel Blackwood, and Peter Kinderman. Persecutory delusions: a review and theoretical integration. *Clinical psychology review*, 21(8):1143–1192, 2001.

Alexander Bird. Dispositions and antidotes. *The Philosophical Quarterly*, 48(191):227–234, 1998.

Paul A Boghossian. The normativity of content. *Philosophical Issues*, 13:31–45, 2003.

Lisa Bortolotti. *Delusions and Other Irrational Beliefs.* Oxford University Press, 2009.

Robert Brandom. *Making it explicit: Reasoning, representing, and discursive commitment.* Harvard university press, 1994.

Jason Brennan. *Against democracy.* Princeton University Press, 2016.

Lara Buchak. Belief, credence, and norms. *Philosophical studies*, 169:285–311, 2014.

Peter Carruthers. On being simple minded. *American Philosophical Quarterly*, 41(3):205–220, 2004.

Clark A Chinn and William F Brewer. The role of anomalous data in knowledge acquisition: A theoretical framework and implications for science instruction. *Review of educational research*, 63(1):1–49, 1993.

Annalisa Coliva. Which hinge epistemology? *International Journal for the Study of Skepticism*, 6(2-3):79–96, 2016. doi: 10.1163/22105700-00603002.

Joel Cooper. *Cognitive dissonance: 50 years of a classic theory.* Sage, 2007.

Edward Craig. *Knowledge and the state of nature: An essay in conceptual synthesis.* Oxford University Press, 1991.

Jonathan Dancy. Knowledge and the state of nature: An essay in conceptual synthesis. *Philosophical Quarterly*, 42(168):393–395, 1992. doi: 10.2307/2219696.

Donald Davidson. Rational animals. *dialectica*, 36(4):317–327, 1982.

Donald Davidson. *Problems of rationality*. Oxford University Press, 2004.

Daniel C. Dennett. True believers: The intentional strategy and why it works. In A. F. Heath, editor, *Scientific Explanation: Papers Based on Herbert Spencer Lectures Given in the University of Oxford*, pages 150–167. Clarendon Press, 1981.

Kevin Dorst. Rational polarization. *Philosophical Review*, 2023.

Hadyn D Ellis and Michael B Lewis. Capgras delusion: a window on face recognition. *Trends in cognitive sciences*, 5(4):149–156, 2001.

Carolina Flores. Delusional evidence-responsiveness. *Synthese*, pages 6299–6230, 2021. doi: 10.1007/s11229-021-03070-2.

Carolina Flores. Resistant beliefs, responsive believers. ms.

Carolina Flores and Elise Woodard. Epistemic norms on evidence-gathering. *Philosophical Studies*, forthcoming.

Daniel Freeman, PA Garety, and E Kuipers. Persecutory delusions: developing the understanding of belief maintenance and emotional distress. *Psychological medicine*, 31(7):1293, 2001.

Jane Friedman. Inquiry and belief. *Noûs*, 53(2):296–315, 2019.

Daniel Gilbert. *Stumbling on happiness*. Knopf, 2006.

James L Gould, Carol Grant Gould, et al. *The honey bee*. Scientific American Library, 1988.

Michael Hannon. *What's the point of knowledge?: a function-first epistemology*. Oxford University Press, 2018.

Michael Hannon. Disagreement or badmouthing? the role of expressive discourse in politics. In Elizabeth Edenberg and Michael Hannon, editors, *Political Epistemology*. Oxford: Oxford University Press, 2021.

Larisa Heiphetz, Casey Lee Landers, and Neil Van Leeuwen. Does think mean the same thing as believe? Linguistic insights into religious cognition. *Psychology of Religion and Spirituality*, 13(3):287, 2021.

Grace Helton. If you can't change what you believe, you don't believe it. *Noûs*, 2020. doi: 10.1111/nous.12265.

Thomas Kelly. Disagreement, dogmatism, and belief polarization. *Journal of Philosophy*, 105(10):611–633, 2008. doi: 10.5840/jphil20081051024.

Ziva Kunda. The case for motivated reasoning. *Psychological bulletin*, 108(3):480, 1990.

Neil Van Leeuwen. Religious credence is not factual belief. *Cognition*, 133(3):698–715, 2014. doi: 10.1016/j.cognition.2014.08.015.

Hannes Leitgeb. The stability theory of belief. *The Philosophical Review*, 123(2):131–171, 2014.

Jennifer S Lerner, Ye Li, Piercarlo Valdesolo, and Karim S Kassam. Emotion and decision making. *Annual review of psychology*, 66(1), 2015.

Alan M Leslie. Pretense and representation: The origins of "theory of mind.". *Psychological review*, 94(4):412, 1987.

Falk Lieder and Thomas L Griffiths. Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences*, 43, 2020.

Tania Lincoln and Emmanuelle Peters. A systematic review and discussion of symptom specific cognitive behavioural approaches to delusions and hallucinations. *Schizophrenia Research*, 203:66–79, 01 2018. doi: 10.1016/j.schres.2017.12.014.

Eric Mandelbaum. Troubles with Bayesianism: An introduction to the psychological immune system. *Mind and Language*, 34(2):141–157, 2019. doi: 10.1111/mila.12205.

Gregory B Markus. Stability and change in political attitudes: Observed, recalled, and "explained". *Political behavior*, 8(1):21–44, 1986.

John McDowell. *Mind and world*. Harvard University Press, 1996.

Victoria McGeer and Philip Pettit. The self-regulating mind. *Language & Communication*, 22(3):281–299, 2002.

Conor McHugh. Epistemic responsibility and doxastic agency. *Philosophical Issues*, 23:132–157, 2013.

Conor McHugh and Daniel Whiting. The normativity of belief. *Analysis*, 74(4):698–713, 2014.

Ruth Garrett Millikan. Thoughts without laws; cognitive science with content. *The Philosophical Review*, 95(1):47–80, 1986.

Steffen Moritz, Christina Andreou, Brooke C Schneider, Charlotte E Wittekind, Mahesh Menon, Ryan P Balzan, and Todd S Woodward. Sowing the seeds of doubt: a narrative review on metacognitive training in schizophrenia. *Clinical psychology review*, 34(4):358–366, 2014.

C Thi Nguyen. The seductions of clarity. *Royal Institute of Philosophy Supplements*, 89:227–255, 2021.

C. Thi Nguyen. Hostile epistemology. *Social Philosophy Today*, 39:9–32, 2023. doi: 10.5840/socphiltoday2023391.

Josef Perner. *Understanding the representational mind*. The MIT Press, 1991.

Jake Quilty-Dunn and Eric Mandelbaum. Against dispositionalism: belief in cognitive science. *Philosophical Studies*, 175(9):2353–2372, 2018.

Susanna Schellenberg. *The unity of perception: Content, consciousness, evidence.* Oxford University Press, 2018.

Laura Schroeter and François Schroeter. Rationalizing self-interpretation. In *The Palgrave handbook of philosophical methods*, pages 419–447. Springer, 2015.

Eric Schwitzgebel. In-between believing. *The Philosophical Quarterly*, 51(202):76–82, 2001.

Eric Schwitzgebel. The pragmatic metaphysics of belief. In Cristina Borgoni, Dirk Kindermann, and Andrea Onofri, editors, *The Fragmented Mind.* Oxford University Press, 2021.

Wilfrid Sellars. Empiricism and the philosophy of mind. *Minnesota studies in the philosophy of science*, 1(19):253–329, 1956.

Declan Smithies, Preston Lennon, and Richard Samuels. Delusions and madmen: against rationality constraints on belief. *Synthese*, 200(3):223, 2022.

JV Stone, IS Kerrigan, and J Porrill. Where is the light? Bayesian perceptual priors for lighting direction. *Proceedings of the Royal Society B: Biological Sciences*, 276(1663): 1797–1804, 2009.

Peter Strawson. Freedom and resentment. In *Proceedings of the British Academy, Volume 48: 1962*, pages 1–25. 1962.

Ema Sullivan-Bissett. Biological function and epistemic normativity. *Philosophical Explorations*, 20(1):94–110, 2017. doi: 10.1080/13869795.2017.1287296.

Joshua B Tenenbaum, Charles Kemp, Thomas L Griffiths, and Noah D Goodman. How to grow a mind: Statistics, structure, and abstraction. *science*, 331(6022):1279–1285, 2011.

Maura Tumulty. Managing mismatch between belief and behavior. *Pacific philosophical quarterly*, 95(3):261–292, 2014.

David Velleman. On the aim of belief. In David Velleman, editor, *The Possibility of Practical Reason*, pages 244–81. Oxford University Press, 2000.

Ralph Wedgwood. The aim of belief. *Philosophical perspectives*, 16:267–297, 2002.

Evan Westra. Symbolic belief in social cognition. *Philosophical Perspectives*, forthcoming.

Aaron Z. Zimmerman. *Belief: A Pragmatic Picture.* Oxford University Press, 2018.