NATURALNESS IS NOT AN AIM OF BELIEF

INTRODUCTION

Recently several authors have defended the thesis that naturalness, in the sense of Lewis (1983), is an aim of belief.[1] This thesis is thought to be motivated by cases like the following. Imagine a community of language users that used the predicates 'bred' and 'rue' that cut across the interpretations we give to 'red' and 'blue'. Ted Sider (2011, p. 2) thinks that in cases like this, "It is almost irresistible to describe these people as *making a mistake.*" Speaking the bred and rue language makes one apt for criticism along some distinctively epistemic dimension. When one believes that something is grue, they make a mistake, they believe incorrectly.

This example can be coherently developed so that the speakers of this language lack false beliefs about what is bred and rue, lack beliefs not amounting to knowledge about what is bred and rue, and lack beliefs not supported by their evidence about what is bred and rue. Whatever mistake they make by theorizing in terms of these notions, it is not a matter of believing falsely, without knowledge, nor without proper support. The mistake is supposed to be one of categorization: users of the 'bred' and 'rue' language have beliefs that fail to match up to the structure of the world because they categorize things in terms of the non-joint carving properties of bred and rue, which draw arbitrary distinctions, instead of the properties red and blue, which draw more natural, or less arbitrary, distinctions. They make a mistake *simply because* they have beliefs who contents ascribe a gerrymandered property. Naturalness, so the argument goes, is an aim of belief, akin to the more recognized aim of truth.

---

[1] See for instance Sider (2011) and McDaniel (2017). Dasgupta (2018) further argues that while this thesis is not often explicitly endorsed, it is an implicit commitment of many of those who theorize about naturalness.

While there is certainly some plausibility to this thesis, I am going to argue in this paper that it is wrong. Naturalness is not an aim of belief. My argument will turn on certain implausible consequences concerning our beliefs about, and inquiry into, naturalness itself, that result from taking naturalness as (one of) belief's aims. If the arguments of this paper are correct, we should resist the inclination to say that one is making a mistake in theorizing in terms of bred and rue as opposed to red and blue. Of course, in theorizing in terms of bred and rue *instead* of red and blue, one thereby fails to know some interesting facts about the colors red and blue. This may be regrettable. But on my view it is no more of a mistake than choosing to study history instead of physics: one doesn't inherently fail to live up to some standard of correctness on belief by doing so, though one does inevitably miss out on some interesting facts about the world. Of course, if one is antecedently interested in finding the natural joints in the world, then one should seek out those joints. But this obligation is hypothetical, not arising from any constitutive aim of belief.

Here is an overview of what is to come. In section 1, I distinguish several different interpretations of the slogan "naturalness is an aim of belief" using the more familiar idea that truth is an aim of belief as my guide. In section 2, I formulate a plausible, and quite standard, principle that connects the naturalness of properties to the naturalness of propositions–the contents of belief–and argue that this principle, when combined with the thesis that belief aims at naturalness, has implausible consequences. In section 3, I further support one of the main premises of the argument by showing that possible replacements for it fail to provide a robust enough theory. I end in section 4 with some more broad brush remarks on what the takeaways of the argument are for our theorizing about naturalness. Here in particular I argue, contra Sider (2011) and Dasgupta (2018), that the interest in naturalness is not dependent on its serving as an aim of belief, and so despite my arguments, naturalness can still be viewed as an important metaphysical posit.

1. Naturalness and the Aim of Belief

My aim in this section is to clarify our subject matter by clarifying the concept of an aim of belief. I'll begin by looking at some of the main accounts of what it is for belief to aim at truth. I then use these accounts as a guide in developing several different accounts of what it is for belief to aim at naturalness. In the next section we'll see that all of these formulations have a common problem, though some to a greater extant than others.

1.1. **Truth as an Aim of Belief.** The slogan "belief aims at truth" suggests a teleological connection between belief and truth. If believing has an aim or a function, it is one fulfilled when the things believed are true. But how to more precisely spell out this idea is far from clear. I will focus on three different interpretations of the slogan: one in terms of *correctness*, one in terms of *obligation*, and one in terms of *value*. I choose these three because they are the concepts that those who advocate for the view that belief aims at *naturalness* use to articulate their theses.[2] There may be interpretations of the slogan using other concepts. And these interpretations may fare better with respect to the argument I will give. But the objections raised in this paper seem to me to challenge every formulation of the claim that has been defended in the literature on naturalness, and so the choice to criticize these formulations is apt.

The interpretation of the slogan that appeals to standards of correctness can be minimally formulated as follows:

CT: One correctly believes p only if p is true.

The thesis CT is a thesis about when beliefs "get things right." If one believes that it is raining while it isn't, they are not getting things right; they are believing *incorrectly*. Note that CT has a reading on which it is perhaps a trivial truth. On one interpretation of 'correct' it means 'accurate'. On that reading CT simply states that one accurately believes

---

[2]Both Sider (2011) and Dasgupta (2018) use all three concepts, not clearly distinguishing between the various different formulations. McDaniel (2017) more explicitly chooses to formulate the thesis in terms of obligation, but notes that he is not wedded to that notion, and that other concepts may work equally well in formulating the theory.

p only if p is true. This isn't clearly a substantive hypothesis since, plausibly, what it is for one to accurately believe p is for one to believe p while p is true. On the intended interpretation of CT, it is a *normative* hypothesis, making a statement about the standards of correctness for belief, akin to the standards of correctness of a game.

Arguably, this normative reading of CT can be brought out more clearly by explaining it in terms of a constitutive *norm* that bans the having of false beliefs:

> NT: One must: believe p only if p is true.

The thesis NT is about our epistemic obligations. If one believes that it is raining while it isn't, they are doing something epistemically impermissible; they make a mistake *by* doing something they should not be doing.

The relationship between NT and CT is contentious. One may accept the former but not the latter, for instance. For my purposes it suffices to note that both have been put forward as possible interpretations of the slogan and so serve as examples to look to in attempting to interpret the the thesis that belief aims at naturalness.[3]

The third interpretation of what it means to say that belief aims at truth is that true beliefs are *intrinsically epistemically valuable.* Explicitly:

> VT: If p is true, it is intrinsically epistemically valuable to believe p.

The thesis VT is not about the possible good effects of having true beliefs. It asserts that true beliefs are good regardless of their effects. They are good in themselves. To believe that it is raining while it is raining is intrinsically valuable. True beliefs are good beliefs.

In the literature on aims of belief, there is disagreement as to whether VT should be counted as an interpretation of the slogan.[4] I do not want to take side on this issue. It is important to note, however, that the value theoretic claim is perhaps the most popular formulation of the idea that *naturalness* is an aim of belief. For instance, Sider states that

---

[3]Normative interpretations of the claim that belief aims at truth are widespread, with some preferring to interpret it in terms of correctness and others in terms of obligation. For a partial list of authors discussing this see Boghossian (2003), Brandom (1994), Engel (2007), Humberstone (1992), Wedgewood (2002), and Williamson (2000).

[4]For some relevant discussion see Lynch (2004) and Williams (2002).

it's *better* to think and speak in joint-carving terms. We ought not speak

the 'grue' language, nor think the thoughts expressed by its simple sentences.

(2011, p. 72)

And follows this up with:

[Natural] thought does not have merely instrumental value. It is rather a

constitutive aim of the practice of forming beliefs, as constitutive as the more

commonly recognized aim of truth (2011, p. 72)

As Sider sees things, notions of value, obligation and aims are all intimately connected. Dasgupta (2018) also seems to take the idea that naturalness is a standard of correctness on belief to be intimately connected with this value theoretic claim. For instance he states that one possible gloss on the claim that naturalness is a standard of correctness on belief can be glossed as the idea "that [natural] beliefs are *better* than [non-natural] ones" (2018, p. 291). Whether or not correctness conditions should be glossed in terms of value, both the claim of value and the claim of correctness are being made by those discussing the idea that naturalness is an aim of belief and so it is important that they both be included.

1.2. **Naturalness as an Aim of Belief.** With these interpretations of the notion that truth is an aim of belief in hand, we can now look more closely at what it might mean to say that *naturalness* is an aim of belief. An obvious strategy is to simply substitute in the notion of a perfectly natural proposition for the notion of a true proposition in each of the above theses:

CN: One correctly believes p only if p is perfectly natural.

NN: One should: believe p only if p is perfectly natural.

VN: If p is perfectly natural, then it is intrinsically valuable to believe p

Do these serve as good interpretations of the idea that belief aims at naturalness? There are two problems that immediately arise, one more important than the other. The first worry is just that we haven't yet said how to extend the notion of perfect naturalness from properties to propositions. In the next section I sketch a theory of how to do this that seems

to me to represent a pretty standard solution in the literature. There is a more substantive worry with these theses however. The problem is that none of them account for the initial data that was supposed to motivate our taking naturalness as one of belief's aims. Neither the proposition that something is bred nor the proposition that something is red is perfectly natural, since neither the property of being bred nor the property of being red is perfectly natural. So by CN's lights, neither bred nor red thoughts are correct. By NN's lights, both bred and red thoughts are impermissible. And for all that VN says, neither bred thoughts nor red thoughts are valuable. According to all of these theories, the mistake in believing that something is bred is similarly made by believing that something is red.

One salient difference between naturalness and truth is that whereas propositions can be more or less natural, they cannot be more or less true. Our interpretation of the slogan that belief aims at naturalness should take this into account. The *relative* naturalness of the properties involved ought to matter when deciding how to believe. It is *more* correct, something that we have *more* reason to do, or *better* to theorize in terms of 'red' as opposed to 'bred' since being red is *more* natural than being bred, even if neither is perfectly natural.

Some care is called for here, however. For instance, perhaps it is better to *truly* believe that something is red than to *truly* believe that something is bred, but *worse* to falsely believe that something is red than to falsely believe that something is bred. Getting things wrong with respect to arbitrary distinctions isn't so bad; getting things wrong with respect to natural ones is. I suggest the following theses as at least partial elucidations of our slogan:

> CRN: If p and q are both true, and p is more natural than q, then it is more correct to believe p than it is to believe q.

> NRN: If p and q are both true, and p is more natural than q, then one has greater prima facie obligation to believe p than to believe q.

> VRN: If p and q are both true, and p is more natural than q, then it is prima facie better to believe p than to believe q.

We can think of a belief as being more correct than another as a matter of getting closer to satisfying the normative standards governing belief than another. Thus the idea behind CRN is that in truly believing that something is red, for instance, one gets closer to "getting things right" than one does when they truly believing that something is bred.

The thesis NRN is closely related to the thesis put forward by McDaniel (2017). It is important here that we restrict ourselves to "prima facie obligations" since there are going to be cases in which an agent has much more evidence supporting some proposition p than they have supporting another proposition q, despite p's being less natural than q. We want to allow in such cases that all things told, they are more obligated to believe p than to believe q.

The thesis VRN is quite close to the formulation preferred by Sider (2011). Again I have qualified the claim so that the claim that believing p is better than believing q is only *prima facie* true when p is more natural than q. Perhaps there are other features that contribute to the value of a belief that ultimately allow beliefs with less natural contents to be more valuable than beliefs with more natural contents.

In the next section, I will argue that none of CRN, NRN nor VRN should be accepted. Since these theses seems to me to exhaust the options for what those who have advocated for naturalness as an aim of belief have meant, the conclusion of my argument is that naturalness is not an aim of belief, at least not in the sense that Sider and others have intended.

2. Naturalness is not an aim of belief

The argument of this section turns on a general thesis that relates the naturalness of properties to the naturalness of propositions. The need for some sort of bridge is clear: while it is primarily properties to which naturalness is ascribed, it is propositions that are believed. If we are to take naturalness as an aim of belief, we thus need some way of connecting the naturalness of properties to the things that we believe.

In the literature, this bridge is often made by tacit appeal to the structured theory of propositions.[5] If we view propositions as structured, having properties and individuals as constituents, then we can attempt to gauge the relative naturalness of a proposition in terms of the relative naturalness of its constituents.

The structured theory of propositions, however, is controversial. It would good if our theory of the naturalness of propositions didn't have to assume it. We could always simply take propositional naturalness as a primitive. But even if we do this, we should expect some general laws connecting naturalness as it applies to properties to naturalness as it applies to propositions that allow us to infer, for instance, that the proposition that something is blue is not perfectly natural from the premise that the property of being blue is not perfectly natural.

2.1. **Type Theory.** To solve this problem I will follow Dorr and Hawthorne (2013) in formulating claims about naturalness in the language of higher-order logic. In this framework, a collection of syntactic categories is inductively defined. There is a type e for names, and for any types $\tau_1, \ldots, \tau_n$, a type $\langle \tau_1, \ldots, \tau_n \rangle$ for expressions that combine with expressions of type $\tau_1, \ldots, \tau_n$ to form sentences. Thus sentences themselves have type $\langle \rangle$, monadic predicates like 'red' and 'bred' have type $\langle e \rangle$, connectives like negation $\neg$ have type $\langle \langle \rangle \rangle$, and so on.

To formulate claims about relative naturalness of various types we can use a typed collection of *higher-order predicates*. For any type τ we suppose there is a predicate $\mathrm{MoreNatural}_\tau$ of type $\langle \tau, \tau \rangle$ in terms of which various comparisons of naturalness can be stated. So for example the sentence 'Redness is more natural than bredness' can be formalized using the higher-order predicate $\mathrm{MoreNatural}_{\langle e \rangle}$ as follows:

$$\mathrm{MoreNatural}_{\langle e \rangle}(\mathrm{Red}, \mathrm{Bred})$$

---

[5]This appeal is explicit in McDaniel (2017) and seems to me implied in Dasgupta (2018). Sider (2011) appeals instead to the concepts needed to "simply express" the proposition in question. The success of this strategy of course turns on what simply expressing amounts to.

Here we have a second level predicate combining directly with two first-level predicates, rather than a first level predicate combining with two names (e.g., 'redness' and 'bredness').

The benefit of this approach is the following: since we have comparative naturalness predicates that combine directly with both predicate and sentences, this allows us to state certain bridge principles connecting them schematically in terms of a well defined notion of substitution on sentences.[6] In particular, the following principle seems to capture many of our intuitions regarding the relationship between the naturalness of properties and the naturalness of propositions:

BP: $\text{MoreNatural}_{\langle e \rangle}(F, G) \rightarrow \text{MoreNatural}_{\langle \rangle}(\varphi[G//F], \varphi)$

The principle BP is a schema. An instance of this schema is obtained by replacing F and G with first-order predicates p and q, replacing φ with a sentence s *in which p occurs*, and replacing $\varphi[G//F]$ with a sentence that results from replacing *all* occurrences of p with q (excepting those that occur within quotation marks). So for example an instance of BP might tell us that the proposition that something is red is more natural than the proposition that something is bred via the following instance:

$$\text{MoreNatural}_{\langle e \rangle}(\text{Red}, \text{Bred}) \rightarrow \text{MoreNatural}_{\langle \rangle}(\exists x \text{Red}(x), \exists x \text{Bred}(x))$$

The requirement that p occur in the sentence with which φ is replaced allows us to avoid counting as instances of BP false claims like:

$$\text{MoreNatural}_{\langle e \rangle}(\text{Red}, \text{Bred}) \rightarrow \text{MoreNatural}_{\langle \rangle}(\exists \text{Blue}(x), \exists x \text{Blue}(x))$$

The proposition that something is blue is not more natural than the proposition that something is blue, despite the property of being red being more natural than the property of being bred. But the sentence that results from '$\exists x \text{Blue}(x)$' by substituting all occurrences of 'Bred' with 'Red' just is '$\exists x \text{Blue}(x)$.' Given our definition of what it is to be an instance of BP, these sorts of problem cases are avoided.

---

[6]See Dorr (2019) for further motivations for adopting the higher-order approach.

The requirement that we substitute *all* occurrences of G with F avoids potential counterexamples to these sorts of bridge principles that have been raised by Dorr (2019). One might think that despite the fact that red is more natural than bred, the proposition that everything bred is bred is more natural than the proposition that everything bred is red:

$$\text{MoreNatural}_{\langle\rangle}(\forall x(\text{Bred}(x) \rightarrow \text{Bred}(x)), \forall x(\text{Bred}(x) \rightarrow \text{Red}(x)))$$

The intuitive thought is just that the naturalness of a given proposition might depend not solely on the naturalness of its constituents, but also on the "homogeneity" of its constituents. The proposition that everything bred is red counts as less natural because is constituents are more diverse than the proposition that everything bred is bred. However all that follows from BP is that the proposition that everything red is red is more natural than the proposition that everything bred is bred, which is intuitively the correct result.

It is worth flagging here though that the principle is not completely neutral on questions of fineness of grain. Since being bred is more natural than being red, it follows that the proposition that something is either red or not red is more natural than the proposition that something is bred or not bred. On a coarse grained view of propositions, these propositions might be identified. But nothing can be more natural than itself.

This is indeed a worry. But it seems to me not too surprising that a theory of relative naturalness will be incompatible with a coarse grained view of propositions. A large part of the push towards "post modal metaphysics," to use Sider's phrase, is the push to use concepts such as grounding, priority and essence which draw more fine grained distinctions than are compatible with, for instance, the possible worlds theory of propositions. [7]

The idea of extending naturalness to items beyond the predicate requires similar more fine grained theories. Indeed many of the intuitive claims we might want to make about naturalness are not compatible with a coarse grained view of propositions, properties and relations. For instance, intuitively the property of being an electron is more natural than

---

[7]See Sider (2020).

the property of being an electron and either a cow or a non-cow. On a coarse grained view of properties, these turn out to be one and the same property.

2.2. **Naturalness is not an Aim of Belief.** Let $<_B$ be a predicate of type $\langle\langle\rangle, \langle\rangle\rangle$. Thus $<_B$ combines with two formulas to make a formula. Intuitively it corresponds to a relation between propositions. We can pronounce '$\varphi >_B \psi$' as 'It is better to believe that φ than it is to believe that ψ'. However officially we treat $<_B$ as schematic. Different principles will result depending on one's preferred way of capturing the idea that naturalness is an aim of belief. For instance, if one preferred the gloss in terms of correctness, then $\varphi >_B \psi$ could be read as "It is more correct to believe that φ than it is to believe that ψ."

The principles CRN, NRN and VRN all have the same form. This allows us to view each of them as a possible interpretation of the following schema:

NAB: $\text{MoreNatural}_{\langle\rangle}(\varphi, \psi) \to (\varphi \wedge \psi \to \varphi >_B \psi)$

This is a schema instances of which result from replacing φ and ψ by formulas. Intuitively, it says that if the proposition that φ is more natural than the proposition that ψ, it is better to believe that φ than to believe that ψ, provided both propositions are true. If we interpret $>_B$ in terms of correctness, this just formalizes CRN. If we interpret $>_B$ in terms of prima facie obligation, it formalizes NRN; and if interpreted in terms of prima facie value, it formalizes VRN.

The conjunction of BP with NAP seems to me to have objectionable consequences. The consequences turn on our inquiry into the notion of perfect naturalness. For each type τ, let $\text{PerfectlyNatural}_\tau$ be a predicate of type $\langle\tau\rangle$ that predicates perfect naturalness of entities of type τ. Then every instance of the following schema can be derived from BP

$$\text{MoreNatural}_{\langle e\rangle}(F, G) \to \text{MoreNatural}_{\langle\rangle}(\neg\text{PerfectlyNatural}(F), \neg\text{PerfectlyNatural}(G)) \quad (1)$$

That is, if a property F is more natural than a property G, then the proposition that F is *not* perfectly natural is more natural than the proposition that G is not perfectly natural. Let '$\neg\text{PerfectlyNatural}(F, G)$' abbreviate '$\neg\text{PerfectlyNatural}(F) \wedge \neg\text{PerfectlyNatural}(G)$'. Then

from (1) and NAB we can derive every instance of:

$$\big(\text{MoreNatural}_{\langle e\rangle}(F, G) \wedge \neg\text{PerfectlyNatural}(F, G)\big) \rightarrow$$
$$(\neg\text{PerfectlyNatural}(F) >_B \neg\text{PerfectlyNatural}(G)) \tag{2}$$

That is, if the property F is more natural than the property G, and F and G are indeed not perfectly natural, then it is better to believe that F is not perfectly natural than it is to believe that G is not perfectly natural. In brief, the closer a property comes to being perfectly natural, the better it is to believe that it is *not* perfectly natural, provided it is indeed not perfectly natural. This claim, it seems to me, is false.

Now as I mentioned, the operator $>_B$ is here being used schematically. We get different interpretations of (2) depending on which interpretation of the claim that belief aims at naturalness we prefer. I think the claim is false no matter how it is interpreted, though some interpretations may be more defensible than others. I'm going to briefly run through what (2) says on these various different interpretations before providing my case that (2) ought to be rejected.

On the correctness-theoretic interpretation, (2) states that when the property F is more natural than the property G, and neither of them are perfectly natural, it is more correct to believe that F is not perfectly natural than it is to believe that G is not perfectly natural. So for example, since blue is more natural than rue, and both fail to be perfectly natural, it is more correct to believe that blue is not perfectly natural than it is to believe that rue is not perfectly natural. The closer a property gets to being perfectly natural, the more correct it is to believe that it is not perfectly natural, provided it is indeed not perfectly natural.

On deontic formulation, (2) states that whenever F is more natural than G, and neither of them are perfectly natural, we are *prima facie* more obligated to believe that F is not perfectly natural than we are to believe that G is not perfectly natural. The closer a property gets to being perfectly natural, the greater our prima facie obligation to believe that it is not perfectly natural.

On the value theoretic formulation, (2) states that whenever F is more natural than G, and neither of them are perfectly natural, it is prima facie better to believe that F is not perfectly natural than it is to believe that G is not perfectly natural. As properties get closer to being perfectly natural, the better it is to believe that they are not perfectly natural.

I think each of these specific instances should be rejected. In the remainder of this section, I am going to defend this view. Since (2) follows from the conjunction of BP and NAB, this means that at least one of those schemas has a false instance. Since BP seems to me to represent the sort of standard theory connecting the naturalness of properties to the naturalness of propositions, this suggests that it is NAB to blame: naturalness is not an aim of belief.

The most straightforward reason to reject (2) is that it is unintuitive. For example, suppose we substitute in 'red' for F and 'bred' for G. On its value theoretic interpretation, it entails that it is prima facie better to believe that red is not perfectly natural than it is to believe that bred is not perfectly natural. On the deontic formulation, it entails that we are prima facie more obligated to believe that red is not perfectly natural than we are to believe that bred is not perfectly natural. And finally on its correctness-theoretic interpretation it entails that it is more correct to believe that red is not perfectly natural than it is to believe that bred is not perfectly natural. None of these claims strike me as particularly intuitive.

Take the correctness-theoretic interpretation. I have a hard time hearing a true reading of "the belief that red is not perfectly natural is more correct than the belief that bred is not perfectly natural." Of course one could say that the belief that red is not perfectly natural is more correct because it is about more natural notions than the belief that bred is not perfectly natural. But this is just to re-assert the theory. If the theory is to be substantive, it must be formulated in terms some notion of correctness on which we have a prior grip. That notion was supposed to be intimately connected to our intuitions about when beliefs "get things right." And it just strikes me as implausible that the belief that red is not

perfectly natural is closer to getting things right than the belief that bred is not perfectly natural.[8]

Suppose that we instead interpret the claim in terms of prima facie obligation. Is the belief that red is not perfectly natural one we are more prima facie obligated to have than the belief that bred is not perfectly natural? One might defend this view by saying that the prima facie obligation exists, but is outweighed by other sorts of obligations. For example perhaps there are obligations to be informed about the field of inquiry one is engaged in. If our field of inquiry is *naturalness*, being informed means knowing that bred is not perfectly natural: properties like bred, grue and rue serve as anti-paradigms for naturalness and so it is important to recognize them as such. This is compatible with their being a prima facie obligation to believe that red is not perfectly natural that is weightier than the prima facie obligation to believe that bred is not perfectly natural.

However I think this response brings out a deeper worry with (2): if it is true, then there are conflicts between the aim of belief and certain norms of inquiry *as a matter of necessity*, not merely as a matter of contingent fact. Consider the norm which states that one should always identify the paradigms and anti-paradigms of any notion that one is theorizing in terms of. If we are theorizing about color, we should be able to identify some clear cases of objects that are red, and some clear cases of objects that are blue. This norm is at odds with the claim that naturalness is the aim of belief. For if naturalness is the aim of belief, then *necessarily* it will always be prima facie better or more correct to form the belief that some property close to being perfectly natural is not perfectly natural than to form the belief that

---

[8]An anonymous reviewer suggested one possible response here. One might insist that the various standards of correctness cannot be added together in order to obtain one overall standard of correctness on belief. All we can say is that the belief that red is not perfectly natural does better with respect to naturalness than the belief that bred is not perfectly natural. I do think that this is a possible line of response, but I'm not exactly sure how the details are supposed to go. It is certainly true that the former belief does better with respect to naturalness than the latter. What is needed, however, is a way of spelling out the normative upshot of this claim. If the normative upshot is not that it is more correct, some other account is needed to say what it is. Simply saying that it is less correct according to the rule that more natural beliefs are more correct won't do because it is not to make a normative judgment at all.

some gerrymandered property is not perfectly natural. But the norm of identifying anti-paradigms suggests that we ought to always find some gerrymandered properties to serve as our anti-paradigms of perfect naturalness.

Now the idea that certain norms of inquiry might conflict with the aims of belief is not itself a worry. After all, it is plausibly a norm that one should believe what one's evidence supports despite the fact that one's evidence sometimes supports falsehoods. The difference is that in this case, the conflict is contingent. One can always gather more evidence to remove the conflict. The naturalness norm is not like that. The conflict between the aim and the norm that we ought to identify anti-paradigms is built into the formulation of the aim itself. This strikes me as more objectionable.

Finally, I do not think our formulation of the aim of belief should have the result that achieving that aim necessarily puts one at risk of error. But according to (2), properties that are *just barely* less than perfectly natural are those for which it is best to believe–or at least, those for which it is prima facie best to believe–that they are not perfectly natural. Thus if we want to achieve the aim of belief while investigating naturalness, there should be some pressure to find those properties that are just barely less than perfectly natural as examples of properties that are not perfectly natural. This isn't what we in fact do. We play it safe and choose utterly gerrymandered properties to serve as our examples. By my lights, there is nothing wrong, and quite a bit right, about this practice. This is further evidence that something has gone wrong with (2).

For these reasons, I think we should reject (2). I am not claiming that as a matter of fact it is the less natural properties that it is better to believe are not perfectly natural. My only argument is that (2) is *false* and so that it is not always true that believing that more natural properties are not perfectly natural is better than believing that less natural properties are not perfectly natural. The objection is to the generalization we get in (2), not necessarily to any specific instance of that generalization.

3. Reject BP?

There are three ways to respond to this argument. One could reject the claim that NAB captures the idea that naturalness is an aim of belief, reject BP, or accept (2). The principle NAB seems to me to pretty directly formalize what Sider and others have meant by the slogan "naturalness is an aim of belief" and so I will set that response aside. The two best responses to the argument are to either reject BP or accept (2). I've already provided some argument for why I think we should reject (2). This leaves us with BP.

Principles like BP are pretty standard in the literature on naturalness. But that doesn't mean we shouldn't rethink them in light of this argument. One perhaps questionable feature of the particular instance of BP that the argument depends on is that the properties F and G are not being predicated *of* something, but are rather the target of some higher-order predication. One possible proposal is then to just look at those cases in which the relevant properties are predicated of something:

BP*: $\text{MoreNatural}_{\langle e \rangle}(F, G) \rightarrow \text{MoreNatural}(F(t), G(t))$

An instance of BP* is obtained by replacing F and G with predicates of type $\langle e \rangle$ and replacing t with a singular term–an expression of type e. So for example, from BP* we could infer the following generalization (in general I am taking the instances of a schema to be closed under generalization):

$$\text{MoreNatural}_{\langle e \rangle}(F, G) \rightarrow \forall x \, \text{MoreNatural}_{\langle \rangle}(F(x), G(x))$$

If the property F is more natural than the property G, then for any object x, the proposition that x is F is more natural than the proposition that x is G. On its face this seems like a true generalization about naturalness and so perhaps BP* could serve in place of BP in one's overall theory of the epistemic role of naturalness.

However if we replace BP with BP*, the resulting theory is just too weak to be of much interest. For instance, we are not able to infer from BP* that the proposition that something is blue is more natural than the proposition that something is rue. That is, the following is

not an instance of BP*:

$$\text{MoreNatural}_{\langle e \rangle}(\text{Blue}, \text{Rue}) \rightarrow \text{MoreNatural}_{\langle \rangle}(\exists x \text{Blue}(x), \exists x \text{Rue}(x))$$

But this seems like exactly the sort of statement our bridge principle connecting the naturalness of properties to the naturalness of propositions *should* predict.

I don't see any obvious way to restrict BP so as to avoid the argument, while also getting a principle strong enough to capture all of the specific instances we want out of a bridge principle. Since in the higher-order setting, quantifiers *are* higher-order predicates, an outright ban on instances that have occurrences of higher-order predicates would also result in an objectionably weak theory.

4. Conclusion

If the argument of this paper is correct, naturalness is not an aim of belief. How significant is this fact? Some have thought that without the claim that naturalness is an aim of belief, it loses much of its theoretical interest. Dasgupta (2018) states for instance that

> [W]ithout the value-theoretic upshot, [naturalness] loses much of its import. For along with the set of natural properties, there is also the set of gratural properties and countless other sets besides. There are natural and gratural joints. Without the realist's value theoretic claim, there is nothing objectively better about carving the world at its natural joints than its gratural joints. (2018, p. 308)

I think that it is true that there is nothing objectively better, at least intrinsically, about "carving the worlds at its natural joints than its gratural joints." But I don't think this means it loses its import. For a good deal of its import is simply that it is relevant to things we are interested in.[9] We are interested in reference, laws, fundamentality and induction.

---

[9]This is not to deny that the distinction between natural and nonnatural properties is an objective one. My view is that it is an objective matter whether a property is natural or not and that our interest in this distinction derives, in large part, from our prior interest in similarity, laws of nature, reference and so on. There is of course a lot more to be said here concerning Dasgupta's argument, my main point is to deny

Naturalness is plausibly relevant to those things, and so is an interesting thing to investigate. What's not clear is why we must explain our interest in naturalness in terms of the intrinsic value of natural beliefs. After all, there seem to me to be plenty of cases in which there are two different theories of some subject matter, where one uses less joint carving notions, despite the two theories being equally interesting, equally good, and equally correct. Take standard membership based set theories. The primitive notion of these theories is membership (and perhaps being a set). If the iterative conception of sets is right, membership is plausibly a more joint carving notion than the notion of a function between two sets. But there are other set theories, the elementary theory of the category of sets for example, that take the notion of *function* as primitive.[10] I don't see why we should view those who work within this theory as making any kind of mistake, however. Rather, it is simply another interesting theory about sets: by theorizing in terms of the arguably less natural notion of a function, certain structural properties of sets become more visible.

Now of course, if the *goal* is to determine the fundamental theory of the world, then we would be making a mistake if we proposed a theory that didn't use perfectly natural notions. However if that is not our goal, then I don't see why our theories that use less natural notions, for instance, arguably, the elementary theory of the category of sets, need be any *less* interesting because of it. Naturalness is one feature among many that we have prior interest in. That's enough to give it import. The assumption that it is a constitutive aim of belief isn't needed. If the arguments of this paper are correct, it also isn't true.[11]

References

Boghossian, P.A. 2003. "The Normativity of Content." *Philosophical Issues* 13: 31-45.

Brandom, R. 1994. *Making it Explicit.* Harvard University Press.

---

the premise that the upshot of naturalness depends on its value theoretic upshot, and instead insist that its upshot derives from the role it plays in other phenomena of interest.

[10]See, for instance, Leinster (2014).

Dasgupta, S. 2018. "Realism and the Absence of Value." *Philosophical Review* 127 (3):279-322

Dorr, C. 2013. "Reading Writing the Book of the World." *Philosophy and Phenomenological Research* Vol.87(3) 717-724.

Dorr, Cian, and John Hawthorne. 2013. "Naturalness." In Karen Bennett and Dean W. Zimmerman (eds.), *Oxford Studies in Metaphysics*, volume 8, Oxford: Oxford University Press, pp. 3-77.

Dorr, Cian. 2019. "Natural Properties." *Stanford Encyclopedia of Philosophy.*

Engel, P. 2007. "Belief and Normativity." *Disputatio* 2 (23): 197-202.

Goodman, Nelson. 1955. *Fact, Fiction, and Forecast.* Cambridge, MA: Harvard University Press.

Humberstone, I.L. 1992. "Direction of Fit." *Mind,* 101 (401): 59-83.

Leinster, Tom. 2014. "Rethinking Set Theory." *American Mathematical Monthly,* no. 5, 403-415.

Lewis, David. 1983. "New Work for a Theory of Universals." *Australasian Journal of Philosophy* 61: 343-77.

Lynch, Michael. 2004. *True to Life: Why Truth Matters.* Cambridge: MIT Press.

McDaniel, K. 2017. "Normative Accounts of Fundamentality." *Philosophical Issues,* 27.1: 167-183.

Sider, T. 2011. *Writing the Book of the World.* Oxford University Press.

Sider, T.2020. *The Tools of Metaphysics and the Metaphysics of Science.* Oxford University Press.

Wedgewood, R. 2002. "The Aim of Belief." *Philosophical Perspectives* 16 267-97.

Williams, Bernard. 2002. *Truth and Truthfulness: An Essay in Genealogy.* Princeton: Princeton University Press.

Williamson, T. 2000. *Knowledge and its Limits.* Oxford University Press.