Philosophical Review

The Philosophical Review, Vol. 112, No. 4 (October 2003)

How Truth Governs Belief

Nishi Shah

Introduction

Why, when asking oneself *whether to believe that p*, must one immediately recognize that this question is settled by, and only by, answering the question *whether p is true*? Truth is not an optional end for first-personal doxastic deliberation, providing an *instrumental* or *extrinsic* reason that an agent may take or leave at will. Otherwise there would be an inferential step between discovering the truth with respect to *p* and determining *whether to believe that p*, involving a bridge premise that it is good (in whichever sense of good one likes, moral, prudential, aesthetic, all-things-considered, etc.) to believe the truth with respect to *p*. But there is no such gap between the two questions within the first-personal deliberative perspective; the question *whether to believe that p* seems to collapse into the question *whether p is true*.

Nor can the explanation for the way we seamlessly move from the former question to the latter question be that it is merely *psychologically* impossible for us to fail to believe what we ascertain to be true. It is not as though, in deliberating about *whether to believe that p*, the reason that one focuses on *whether p is the case* is that one has noticed that as a matter of psychological fact one has come to believe whatever one has ascertained to be the case. This would involve an inferential step: "Should I believe that *p*? Well, I'll end up believing that *p* if I ascertain that *p* is true, so I might as well consider whether *p* is true." But there is no such inferential step involved in moving from "Should I believe that *p*?" to "Is *p* true?" Within the perspective of first-personal doxastic deliberation, that is, deliberation about what to believe, one cannot separate the two questions. What I mean by claiming that the two questions cannot be separated is that one cannot settle on an answer to the question *whether to believe that p* without taking oneself to have answered the question *whether p is true*. One can certainly reflect upon one's fallibility and recognize that some of one's beliefs might be false. But so long as one is considering the deliberative question of what to believe, these two questions must be viewed as answered by, and answerable to, the same set of considerations. The seamless shift in focus from belief to truth is not a quirky feature of human psychology, but something that is demanded by the nature of first-personal doxastic deliberation. Follow-

Published by Duke University Press

ing Richard Moran, let us call this feature of doxastic deliberation *transparency*.[1]

This paper is an investigation of the basis of transparency. A possible source of resistance to acknowledging that the phenomenon of transparency is deep enough to warrant investigation is the thought that transparency is merely the by-product of the redundancy associated with the truth predicate. Just as adding that *p* is true to one's assertion of *p* is redundant, adding that one believes that *p* to be true adds nothing to one's assertion that one believes that *p*. But notice that redundancy attaches to any attitude that takes a truth-bearer as its object. For example, to imagine something is to imagine it true, to assume something is to assume it true, to desire something is to desire it true, etc. The phenomenon of transparency that I am interested in explaining arises only in the context of deliberation about what to believe, not in deliberation about what to assume, imagine, or anything else. What I want to explain is why discovering evidence that it is true that snow is white is relevant to determining *whether to believe that snow is white* but is not relevant to determining *whether to imagine or assume that snow is white*. Of course, given redundancy, in having reasons to believe or imagine that snow is white, one thereby has reasons to believe or imagine that it is true that snow is white, but transparency is related to the distinct fact that, from the first-personal deliberative point of view, discovering evidence that snow is white gives one only a reason to believe that snow is white, not a reason to imagine that snow is white. Because transparency indicates a substantive relation between belief and truth, it serves to distinguish belief from other propositional attitudes. This relation thus should not be confused with the trivial grammatical point that one can append the truth predicate to any propositional attitude verb.

I will proceed as follows: First I argue that neither *noncognitivist* accounts of normative epistemic concepts such as *justification* or *warrant*, nor *teleological* accounts of *belief*, deliver adequate illumination of the noncontingent, noninstrumental role of truth in structuring doxastic deliberation. I then suggest an alternative explanation of transparency that makes use of the insight contained in, but misinterpreted by, teleological accounts. I argue that the idea that truth is the standard of correctness for belief, which is the key to explaining transparency, is not correctly interpreted by the teleologist's descriptive claim that belief is causally regulated by truth-regarding considerations, but rather is expressed in the *prescription* to believe that *p* only if *p* is true that frames an agent's deliberation about *whether to believe that p*. My

Published by Duke University Press

hypothesis is that accepting this prescription is one of the conditions for possessing the concept of belief. Unlike the teleologist's hypothesis, this hypothesis accounts for the fact that the normative question *whether to believe that p* gives way to the factual question *whether p is true*. A corollary to this hypothesis is that reasons for an affirmative answer to the question *whether to believe that p* must be considerations that are taken as relevant of the truth of *p*. Thus, my hypothesis also accounts for the essentially evidential character of norms of rational belief, something noncognitivist accounts of rationality leave inexplicable.

Because my proposal is that exercising the concept of belief involves accepting that in some sense one *ought* to believe that *p* only if *p* is true, a meta-normative interpretation of this judgment is required in order to give a complete account.[2] I do not provide such an interpretation here; my purpose in this paper is just to argue that we need such a meta-ethics of **belief.**[3]

Gibbard's Noncognitivism

First let us look at the way that noncognitivism might be thought to explain transparency. I will focus on Allan Gibbard's version of noncognitivism, norm-expressivism, as it applies to judgments of warranted or rational belief. I do this because Gibbard is the only noncognitivist I know who directs his account to explaining normative discourse about *belief* (for example, rather than action). The heart of my criticism of Gibbard's noncognitivism is that it is incomplete, and cannot give a satisfying explanation for the necessary connection between truth-regarding considerations and warranted belief.

Norm-expressivism has its historical roots in the noncognitivist tradition. Many noncognitivists took their inspiration from a common but controversial interpretation of Hume, according to which Hume's fundamental insight was that factual belief is not intrinsically action-guiding, whereas normative judgment is. In order to explain this difference between factual belief and normative judgment, noncognitivists have attempted to analyze normative judgments in terms of expressions of motivation-laden states such as desires, preferences, and emotions. If such judgments themselves express motivation-laden states, then it is obvious how what these judgments express can motivate action. This strategy thus captures Hume's insight that normative judgments seem to have an intimate connection to motives for action. Since these motivational states do not appear to have a representational function, non-

Published by Duke University Press

cognitivists have also tended to conclude, like Hume, that normative judgments do not express factual beliefs, do not attempt to describe a state of affairs, and do not have robust truth conditions.

Norm-expressivism is a novel, sophisticated noncognitivist approach to analyzing normative judgments. The key to this approach is the introduction of the theoretical explanatory concept of a motivational state of mind that Gibbard calls "norm-acceptance." States of norm-acceptance are defined in terms of their role in regulating beliefs, actions, emotions, or whatever else might be the object of normative governance. They bear (quasi-)logical relations to each other, and therefore are shaped by considerations that other motivational-laden states might not be, such as (quasi-)consistency and (quasi-)coherence. Furthermore, they are influenced by discussion, and therefore have a social aspect that other, less linguistically infused motivational states might lack. For example, if I accept a norm that prohibits missing departmental meetings, I will be disposed to attend departmental meetings; my norm will be inconsistent with a norm that permits missing meetings when I don't feel like going; in conversation with friends I can express my acceptance of this norm by way of giving its content as my justification for not going out for drinks, and my friends can press me to give reasons for accepting this norm by offering reasons for accepting a different norm instead, one that permits me to skip meetings for the sake of having drinks with friends.[4]

Gibbard's norm-expressivism is an account of normative judgment generally, and judgments of warranted or rational belief are among the normative judgments to which he applies it. According to Gibbard, to judge that the belief that p is rationally required is to express one's acceptance of norms that require believing p under the circumstances. The state of mind that constitutes accepting a system of norms is a motivational state, one that among other things motivates one to be in the state that is required by those norms.[5] For example, if X thinks that it is rationally required that Samson believe Delilah to be loyal, X accepts norms that, for being in Samson's circumstances, require believing Delilah to be loyal.[6] The state of mind is such that it would dispose X to believe Delilah to be loyal if X were to judge himself to be in Samson's circumstances.

In a sense this is the end of Gibbard's story with respect to rationality in belief. There's no further question on his view as to whether there is a property, rationality, that some beliefs have and some don't. Of course, in coming to accept norms for belief, one may in fact be com-

Published by Duke University Press

mitted to there being a property in virtue of which a belief is rational, but it is not the semantic function of the concept of *rational* belief to refer to a property that some beliefs have and others don't. Rather, the term 'rational' serves to give voice to an attitude permissive of beliefs that are allowed by the norms that one accepts.[7]

Now Gibbard does say that pragmatic considerations fail to bear upon what it is rational to believe, and that the only considerations relevant to rationality are evidential:

> Take the stock example of the man who has evidence his wife is unfaithful. Whether it is still rational for him to believe her faithful—whether such a belief would be warranted—depends on his evidence, and on his evidence alone.[8]
>
> Take again the case of the deceived husband: his evidence may make it rational for him *to believe* his wife unfaithful, but the way the belief would affect his feelings toward her may make it rational for him *to want to believe* her faithful. Rationally *feeling* or *believing* something is distinct from rationally *wanting* to feel or believe it.[9]

There are at least two different interpretations of what Gibbard might mean when he says that only evidence warrants belief. According to the first interpretation, he is making a substantive normative judgment expressing the norms that he accepts for governing belief. According to the second interpretation, Gibbard is making a conceptual claim about the analytic normative connection between **belief** and **evidence**. I will deal with each interpretation in turn, arguing that the first cannot adequately explain why only evidence warrants belief, and that the second both is inconsistent with one of the central features of Gibbard's norm-expressivism and cannot capture the phenomenology of doxastic deliberation.

On the first interpretation, Gibbard's claim that only evidence is relevant to the rationality of belief should be read expressivistically. Gibbard is expressing his acceptance of norms that permit evidence to govern belief and prohibit all considerations other than evidence from governing belief. It would be open for someone else to disagree with him and to claim that only pragmatic considerations justify belief. This person would then be expressing an acceptance of norms that permit pragmatic considerations to govern belief and prohibit all other considerations from governing belief.

On this first interpretation, Gibbard would not be claiming that his pragmatist opponent is incoherent or conceptually confused: the pragmatist is merely expressing his own countervailing norm. It would be

Published by Duke University Press

perfectly coherent for other people to accept norms disposing them to believe whatever they found pragmatic reason for believing, just as we accept norms that dispose us to believe whatever we find alethic reason for believing. Gibbard does say that states of norm acceptance tend toward consistency and consensus. One must try to meet demands for consistency in one's norms and one must be willing (or at least not wholly unwilling) to enter dialogue with an aim toward consensus about one's norms if one is to count as accepting norms.[10] But a version of Voltaire's Dr. Pangloss, who accepts prudential norms that license believing that God exists, that this is the best of all possible worlds, and everything else that he believes, could satisfy these demands; seeing as how such pragmatic norms could be mutually consistent and, if as a result of his dazzling sophistry everyone around Pangloss came to accept these prudential norms, the demands for consensus would be met. This example shows that the logical and social constraints that Gibbard places on the concept of norm-acceptance fail to prevent his account from allowing as a live possibility the existence of agents who coherently accept norms for belief that are purely pragmatic in content. What Gibbard can't explain is why, except possibly in fiction, *there are no such people*.

Well, this isn't exactly right. Gibbard might suggest that as a normative matter, accepting such a position would be crazy in the same way that accepting that counting blades of grass is *the thing to do* would be crazy. Such positions are prima facie implausible, and therefore they are not taught or discussed. Furthermore, if there were a species that held such a position with respect to their beliefs, the world, not being so aligned as to make wishful beliefs reliably successful, would grant them a very short life. Believing that it is safe to pet tigers because their fur is pleasant to touch would lead to the extinction of a species in a hurry. [11]

The problem with this proffered explanation of the absence of pragmatically governed believers is that it allows for the possibility of creatures for whom there is an inferential step between their judgments of truth and their beliefs. These would be creatures that take only evidence to warrant belief, but do so on the grounds that it would be advantageous (for example, utility maximizing) to have true beliefs. They would calibrate their beliefs to their evidence, and therefore not get into the kind of trouble that befalls wishful thinkers. Therefore there is no reason that such creatures should not be allowed as contingent evolutionary possibilities by someone who thinks that it is only for

Published by Duke University Press

substantive normative/pragmatic reasons that pragmatic considerations don't warrant belief. But if such creatures existed, their judgments as to the utility of true beliefs would mediate the bearing that their judgments of truth had on their beliefs. However, such an inferential step is ruled out by transparency, as I pointed out earlier. There is no inferential gap between judgments of truth and belief. Even if it is clearly disadvantageous to accept that only pragmatic considerations warrant beliefs, it cannot be for this reason that there are no creatures that hold such a position.

Let us examine more closely how transparency rules out the possibility of such pragmatists. It would be incoherent from within the perspective of first-personal doxastic deliberation to judge that there is evidence that *p* and yet be completely unmoved to believe that *p*. This is because, due to transparency, considerations that are recognized within doxastic deliberation as relevant to determining the truth with respect to *p* must be immediately seen as relevant to determining *whether to believe that p*. Transparency also blocks the possibility of an internal motivational connection between judgments of what it would be advantageous to believe and belief. What I mean by claiming that there is no *internal motivational connection* is that pragmatic judgments don't *automatically dispose* one toward belief. Judgments of the advantageousness of believing *p* aren't directly relevant to determining the truth value of *p*,[12] and transparency permits only considerations that are directly relevant to determining the truth value of *p* to be immediately belief-guiding. Therefore pragmatic considerations are ineligible to play an immediate role in doxastic deliberation.[13]

The lack of resources in Gibbard's theory to rule out norms *allowing* pragmatic reasons for belief will also prevent the theory from ruling out norms *requiring* pragmatic reasons for belief. If it is coherent to claim that pragmatic considerations can be reasons to believe, it might be supposed that it is coherent to claim that *only* pragmatic considerations can be reasons to believe, and I see no resources in Gibbard's theory to rule this out as a coherent conceptual possibility.[14] Accepting the latter claim would require accepting that it is perfectly intelligible for the cuckolded husband to say, "Yes, it may in fact be true that my wife is cheating on me, but that has absolutely no relevance to what I ought to believe about her faithfulness, since personal happiness is the only thing that matters when it comes to figuring out what to believe about one's spouse." In fact, accepting the coherence of an exclusively pragmatist position would require going even further, allowing that

Published by Duke University Press

someone might coherently hold such an attitude toward all of his beliefs.

Now, it is one thing to claim that truth has no relevance with respect to determining a certain class of beliefs (for example, about someone dear to one); it is quite another to claim that truth has no relevance to determining belief in general. I think that it is clear that we would question whether someone who made such a claim had fully grasped either the concept of belief or the concept of truth. Our handle on the concept of belief comes via its connection to truth, and this connection is an internal, normative one. Someone who claimed that truth never has a role in determining rational belief would be denying that the veracity of a belief provides even a *defeasible, non-overriding* reason. This, I suggest, doesn't make sense. Even if I were wrong that only evidential considerations could warrant belief, it is clearly incoherent to hold that evidential considerations are never relevant to rational belief. But according to the interpretation of Gibbard under consideration, on which he rules out such a claim on normative/pragmatic grounds, he must at least allow that the normative position that such a claim expresses is a coherent conceptual possibility, because the only conceptual constraints on the norms one can accept that Gibbard adduces fail to rule out such a position.

The other interpretation of Gibbard's claim that pragmatic considerations are not reasons to believe, but rather reasons for wanting to believe, is that he is making a *conceptual* point. This interpretation would explain why Gibbard claims that a pragmatic consideration is a reason for wanting to believe rather than merely a poor reason to believe. If it were due merely to his own normative perspective (the system of norms he accepts) that Gibbard rejects pragmatic considerations from governing belief, there would be no reason for him to make the further claim that the pragmatic consideration is a reason for desiring to believe. The purpose of adding this presumably is to point out a possible conceptual confusion that one might fall prey to, conflating reasons to believe with reasons for wanting to believe.

Interpreted as a conceptual claim, what Gibbard says seems to me correct. Suppose that the cuckolded husband, upon being warned of this conceptual conflation, were to reply, "Desiring to believe that my wife is faithful does me no good, since it will not make me feel better. It is only by actually believing in her fidelity that my spirits will improve. So in fact the pragmatic consideration that I am adducing is not a reason for wanting to believe she is faithful, but rather really a reason to

Published by Duke University Press

believe she is faithful." If the husband were to argue in this manner, I think that we would doubt his mastery of the concept of a reason for belief, specifically his mastery of the way that truth serves as an independent standard constraining the character of doxastic reasons.[15] It is not that the husband couldn't take the effects of belief on his feelings to be evidence that it is true that his wife is faithful. For example, he might believe in a benevolent god who always makes true what would make him happy, in which case he might conclude from the fact that his wife's fidelity would make him happy that in fact it is true that she is faithful. But the husband in the example doesn't think that this prudential consideration of the effect of the belief on his happiness is evidence of the truth of the belief, yet thinks the prudential consideration gives him a reason for the belief all the same. And this seems unintelligible.[16]

However, I doubt that Gibbard can avail himself of this more plausible interpretation of his comments, because it conflicts with one of the central motivations of his norm-expressivism. According to Gibbard, the norm-expressivist account he offers is meant to yield a concept of rationality, or warrant, that applies to anything that can be under normative governance, whether it be belief, guilt, action or whatever states of mind normative governance is called for.[17] In part, his motivation for seeking such a general normative concept is to accommodate a deep and puzzling feature of seemingly all normative discourse, the possibility of irresolvable disagreement. The concept of warrant is supposed to provide a framework in which fundamental normative disagreements can be expressed, whether it is with respect to questions of what to believe or what to feel or what to do or whatever.[18] And in order to frame fundamental disagreements, the concept of warrant cannot have a fully specified substantive content attached to it, but must express a thin, "flavorless endorsement."[19]

Of course, to judge that a belief is desirable is to endorse it in *some* sense. However, as we have seen, transparency bars considerations of desirability from being reasons to believe. The point is not that one can't positively evaluate a belief on the basis of judgments of pragmatic advantage, as the cuckholded husband does when he judges it desirable to believe that his wife is faithful, but rather that one can't evaluate a belief as *rational*, in the sense that is relevant to answering the deliberative question *whether to believe that p*, on the basis of such considerations.[20] But because Gibbard restricts himself to a thin sense of 'warrant' as a general term of endorsement, it should be possible, on

Published by Duke University Press

his account, to judge a belief warranted on the basis of the reasons one has for thinking it desirable. Gibbard's account thus does not have a way to distinguish norms of desirability from the type of norms that guide doxastic deliberation. He does not have the resources to distinguish these two different dimensions of endorsement precisely because he wants his concept of warrant to be general enough to capture what is common to all genuinely normative judgments. If Gibbard were to force an evidential constraint into the concept of warrant as such, so that he could capture the evidential constraint on norms of rational belief, then he would lose the ability to account for rationality judgments about cognitive attitudes other than belief, such as judgments of what it is rational to assume for the sake of argument or to imagine for the sake of a fantasy. Norms of rationality for these other cognitive attitudes are not evidentially constrained, and therefore a theorist who wants to use a single normative concept to elucidate the many different varieties of normative judgment, as Gibbard does, cannot evidentially constrain that concept. Given Gibbard's purposes, his concept of warrant cannot accommodate an interpretation of his comments that would allow him to distinguish the endorsement involved in judging a belief rational from that involved in judging it desirable.[21]

In any case, remember that the goal of this paper is first and foremost to explain transparency, not to explain the evidential character of doxastic norms. Even if Gibbard were to give up the ambition of giving a general account of normative judgment, instead tailoring his account of **warrant** specifically to capture our evidentially constrained concept of rational belief and thereby availing himself of the second interpretation of his comments, this move would not explain transparency. This is not to make a criticism of Gibbard's theory, since it may very well be adequate to fulfilling his own (now more limited) explanatory ambitions. My point is that a norm-expressivist account of (evidential) **warrant** isn't adequate to my purpose, which is to explain why in asking oneself *whether to believe that p* one is forced to recognize that this question is answered by, and only by, determining *whether p is true*. To see this, notice that transparency would survive our adopting pragmatic norms (if only we could). For example, even if the cuckholded husband were able to accept pragmatic doxastic norms that licensed believing that his wife is faithful, he would still be unable to conclude deliberation until he had discovered (by his lights) whether it is true that his wife is faithful. He would be in the unfortunate position of accepting doxastic norms that he was unable to implement in his dox-

Published by Duke University Press

astic deliberations. In such a case, transparency would be an obstacle to his ability to normatively regulate his beliefs, and he would have to use indirect, nondeliberative means to ensure that he believes that his wife is faithful. But his acceptance of pragmatic norms would not remove transparency, because transparency is a feature of doxastic deliberation as such. The fact that transparency would not be removed by the adoption of pragmatic norms shows that it doesn't depend upon which doxastic norms we adopt. Therefore a conceptual obstacle to the possibility of accepting pragmatic doxastic norms won't help to explain transparency.

Let me now sum up my discussion of Gibbard's norm-expressivism. For all that Gibbard's account of warrant judgments *actually* says (as opposed to the judgments of warrant that Gibbard makes that do not follow from the meta-normative account of the concept of warrant he gives), it is merely a *contingent* fact that we evidentially restrict our doxastic norms. But transparency obtains for all thinkers who engage in doxastic deliberation, and therefore the internal connection between judgments of evidence and belief holds of any creature to which we are willing to attribute the full-fledged concept of belief. It would be incoherent for any creature, not just us, to claim sincerely to believe that p but also to claim sincerely to be unmoved to change his belief by what he takes to be overwhelming evidence against the truth of p.[22] But there is nothing in Gibbard's official account of normative judgment that yields this result. And even if Gibbard could explain the evidential character of doxastic norms by somehow adding substantive conceptual constraints to his account of **warrant**, a move that I have argued is in tension with his ambition of explaining all normative judgments in terms of one all-purpose normative concept, he would not thereby have explained transparency. All that adding conceptual constraints to the concept of warrant can do is to restrict the doxastic norms we can accept, but transparency is independent of the doxastic norms we do or can accept. The basis for transparency must therefore lie elsewhere.

Standards of Correctness and the Aim of Belief

In this section I will describe an idea we might make use of in explaining the connection between belief and truth that is manifested in the phenomenon of transparency. Then I will describe one very seductive but, as I will argue in the next section, ultimately unsatisfying, way of interpreting this idea that has been presented by David Velleman.

Published by Duke University Press

The explanation I have in mind of why doxastic deliberation has an essentially alethic structure is grounded in the idea that truth provides the *standard of correctness* for belief. In order to see how this idea might help us, we first need to understand what it means to attribute a standard of correctness. Gideon Rosen uses the example of calling some sequence of events a performance of Jingle Bells to illustrate the way that standards of correctness function as norms internal to the kinds to which they apply.[23] According to Rosen, to call an action or bit of behavior a performance of Jingle Bells is to imply that it is correctly done if and only if a certain set of notes is played in a certain sequence. The idiom of correctness and incorrectness brings with it a distinctive type of normativity according to Rosen, one that can be captured neither by the language of value nor by the language of obligation and permission. In order to make people laugh, someone might want to play a silly and offbeat version of Jingle Bells as part of a comedic routine. In such a case playing Jingle Bells correctly is not what he should do, since it would not serve his purpose or plausibly fulfill any moral obligation; nor would a mundane, unsurprising correct performance of the song realize a value, since the only value apt for expression in that context is comedic value. So to claim that a playing of Jingle Bells is correct is neither to endorse it as realizing any sort of value, nor to say that it was played as it ought to be played. Furthermore, according to Rosen judgments of correctness need not play any role in governing deliberation:

> From the point of view of deliberation, judgments about correctness function as mere judgments of fact. Whether the judgment motivates me to act in a certain way depends on whether I have an interest in conforming to the rules. Whatever practical weight the judgment possesses—whether it seems to me to weigh for or against a contemplated act—derives from a separately intelligible desire (or something very much like a desire) to do the correct thing, and not from the judgment of correctness all by itself.[24]

> The simple fact that A would be incorrect does not constitute a reason not to perform it unless there is in the background a substantive synthetic principle to the effect that one has a reason not to do what is incorrect.[25]

Here Rosen is claiming that the mere fact of incorrectness does not generate a reason, which entails that the fact that a belief is incorrect does not, by itself, generate a reason to give up the belief. I propose that we investigate why truth is the standard of correctness for belief, and whether truth, as the standard of correctness *for belief*, has the kind

Published by Duke University Press

of intrinsic deliberative import that Rosen claims standards of correctness in general lack.

So what is it about belief that makes truth its standard of correctness? David Velleman argues that belief has a standard of correctness, truth, in virtue of having a constitutive aim:

> Let me be clearer about the relation between the constitutive aim of belief and the norm that applies to belief in light of that aim. To say that belief aims at the truth is not simply to re-express the norm stipulating that a belief must be true in order to be correct; rather, it is to point out a fact about belief that generates this norm for its correctness.[26]

Velleman is claiming that one can explain how there exists a norm of correctness for belief by turning to the fact that belief aims at the truth. Thinking of truth as the aim of belief typically involves attempting to account for the normative relation between belief and truth along *teleological* lines. If belief aims at the truth, then in some sense the goal of belief is truth. According to this type of view, the normative force that truth has for belief is the normative force that a *goal* has for the entity whose goal it is, and goals can be generated in a number of different ways—intention, intelligent design (which may reduce to intention), and evolutionary selection being the most obvious goal-forming mechanisms. Just as a goal provides the one who has it with motivation to pursue means to its obtainment, so too on this view truth serves as a goal that motivates a believer to strive to attain it. If this were the right way of thinking of belief's standard of correctness, that is, as a goal that one must have if one is to be in the business of forming a belief, then it would appear that the standard of correctness for belief couldn't fail to have motivational efficacy.

According to such a view, reasons for belief possess the instrumental normativity associated with a *telos*. Belief's standard of correctness expresses the measure of success that a belief must achieve in order to obtain its goal. This measure of success then determines which types of considerations constitute reasons for belief; they are whichever considerations lead one in the direction of success in obtaining the goal of belief, truth. Normally the way that considerations point one in the direction of truth is by being indications of truth, that is, by constituting *evidence* (of truth).

There does seem to be something right about the slogan that belief aims at the truth. Belief aims at the truth in the same way that someone building a house aims at providing shelter. In attempting to engage in the activity of house building, one must have the intention of providing

Published by Duke University Press

shelter, whatever other intentions one also has.[27] Analogously, one must be guided by the intention of believing that *p* only if *p* is true if one's activity is to count as inquiry, that is, the intentional activity of forming beliefs. If one came to represent *p* as true as a result of merely being guided by an intention to represent as true propositions that it would be entertaining to represent as true, one's activity wouldn't count as inquiry. Thus, by informing the intentions of the agents engaging in these activities, shelter and truth function as constitutive standards of correctness for the activities of housebuilding and inquiry.

I think this analogy is essentially correct, but we need to be careful in deciphering its upshot. I will proceed as follows: First, I explain how the teleologist interprets the analogy. I focus on teleological accounts because I think this species of descriptivist theory of belief has the best chance of explaining the substantive, necessary connection between truth and belief indicated by the phenomenon of transparency. I argue that the account of belief to which their interpretation of the analogy leads them cannot capture the fact that truth only functions as the standard of correctness in one kind of belief-forming process, doxastic deliberation, or at least it cannot capture this fact without clashing with other central platitudes about belief.

In order to uncover a better way of incorporating the notion of a standard of correctness into an account of **belief**, I will need to go back and explain how I think we ought to understand the analogy between belief and aim-constituted activities. I will argue that a better understanding of the analogy requires us to think of **belief** not as a descriptive concept, as teleological accounts would have us do, but as a normative concept.

The Teleologist's Dilemma

The teleologist takes the analogy between belief and other aim-constituted activities as telling us something about the metaphysical nature of belief. He thinks that the analogy shows that unless a mental state is under the controlling influence of an intention to accept the truth, or as I shall explain, some surrogate of such an intention, the mental state does not count as a belief.[28] The reason that the teleologist needs to make room for some surrogate of an intention to accept the truth is that it is clear that most of our beliefs aren't formed through our intentional attempts to accept the truth. So in order to accommodate the vast majority of cases of belief, the teleologist must allow that cognitive

Published by Duke University Press

mechanisms that don't involve an agent's intentions can count as aiming at truth. One way of doing this is to say that if the mechanisms that produce a cognitive state were designed, either by evolution or learning, to produce true cognitions, then those mechanisms count as being aimed at the truth, and the cognitive states that they produce and regulate count as beliefs.[29]

I will argue that any such teleological account is caught on the horns of a dilemma. On one horn, the teleologist must allow the disposition that constitutes aiming at truth to be so weak as to allow paradigm cases in which beliefs are caused by such non-evidential processes as wishful-thinking, in which case he cannot capture the exclusive role of evidence in one particular type of belief-forming process, reasoning. On the other horn, in order to account for the exclusive role of evidence in reasoning about what to believe, the teleologist must strengthen the disposition that constitutes aiming at truth so that it excludes the influence of non-truth-regarding considerations from such reasoning. However, by strengthening the truth-aimed disposition, the teleologist cannot accommodate the cases of wishful thinking, in which non-evidential factors clearly exercise influence over belief. This dilemma arises because teleological accounts try to reduce the essentially normative conceptual relation that truth has to belief to a descriptive, causal relation. Before I try to substantiate these charges, however, I want to bring out what I take to be plausible about teleological accounts.

It does seem plausible that cognitive states that are causally regulated by truth-tracking mechanisms are good candidates for being in the extension of 'belief'.[30] Truth-regulated cognitive states are well suited for the important role in guiding behavior that we assign to beliefs. States that bear these important causal relations to behavior are suited for this role because they generally provide accurate representations of the organism's environment. Unless a creature's behavior is guided by accurate representations of the location of food, for example, food-gathering behavior is bound to be thwarted. Obviously, the accuracy of these cognitive states is maximized if they are regulated by truth-tracking mechanisms. It thus makes sense that the output role of guiding behavior that we assign to beliefs should be occupied by states that are regulated, on the input side, by truth-tracking mechanisms.

But it also seems clear that we label some cognitive states "beliefs" even though those states were not solely produced or sustained by mechanisms aimed at the truth.[31] I'm thinking here of standard exam-

Published by Duke University Press

ples of cognitive states that are produced by wishful thinking or other processes that are sensitive to practical rather than evidential considerations, but that bear the same causal relations to behavior as paradigmatic beliefs. Our cuckholded husband may not be able to treat his weal and woe as reasons to believe his wife is faithful, but it is all too common for such factors to influence his belief nonetheless. Surely those postmodernists who think that reason never influences belief are wrong, but it is entrenched in common sense that reason does not always have decisive doxastic influence either. Our discourse is replete with accusations that likes and dislikes influence belief either directly or more commonly, indirectly, by blocking an agent from seeing evidence contrary to his preferred views or making him susceptible to a fallacious line of reasoning. Recent debates about the alleged connections between IQ and poverty and sociobiological explanations of human sexual behavior are rife with such accusations on both sides, and while individual charges of prejudice might be disputed, the legitimacy of the general form of such criticisms is taken for granted by all sides.

The teleologist can allow for such cases by accepting that the belief-constituting disposition to be influenced by truth-regarding factors needn't be the only or even the strongest disposition influencing belief.[32] What the teleologist takes to be essential to the nature of belief is just that such a truth-aimed disposition enters into its regulation, not that other dispositions be incapable of also influencing belief.[33] Put this way, in terms of a weak disposition to be influenced by evidence, the teleologist's thesis about the truth-directed nature of belief is quite plausible. While the cuckholded husband's interests may play a decisive role in shaping his beliefs about his wife, we would be hard-pressed to call his cognitions beliefs rather than fantasies if they were shown to be insusceptible to the evidence of his wife's transgressions when it were brought to his attention.[34] In any event, there don't seem to be any clear cases of belief that are immune to the influence of evidence, once we accept that this influence may be rather weak and easily overpowered by the force of other non-evidential factors.

However, this way of accounting for non-evidential influences on belief puts the teleologist on one horn of the dilemma, leaving him unable to explain why one particular belief-forming process, reasoning, is regulated *solely* by evidential considerations. The phenomenology of deliberation that is framed by the question "Should I believe that *p*?" is that only considerations that are relevant to determining the

Published by Duke University Press

truth of *p* can provide reasons that would settle the deliberative question. This point about transparency does not imply that non-evidential factors cannot causally influence belief, or that we cannot recognize that such factors may infect our thinking, but rather it implies that we cannot recognize the relevance of such factors in our deliberative reasoning about what to believe. Because the teleologist allows that considerations other than evidence can influence belief, he has no explanation for why deliberation, being just one of many different belief-forming processes, couldn't be influenced by practical considerations through and through. There is certainly no way for him to bar such considerations from playing a role in deliberation, even if he could explain why evidence must also play a deliberative role.

This horn of the dilemma can be represented as a variant of G. E. Moore's famous open question argument. Now it is certainly controversial what, if anything, Moore's open question argument demonstrated.[35] And I certainly think Moore himself was mistaken to think that his test generated a metaphysical result. Although I don't have space to argue for it here, I do think that his test can be used to generate the presumption that 'good' is not synonymous with any set of solely descriptive terms, so that, in this restricted sense, 'good' cannot be given a descriptive definition. Descriptive definitions fail, as Moore illustrated, because conceptualizing a state of affairs in purely descriptive terms does not close the normative question whether the state of affairs is good. Therefore the concept expressed by 'good' has a normative meaning that descriptive definitions fail to capture.

A variant of Moore's test can be used to show that, as a semantic matter, teleological accounts are inadequate to capture the meaning of 'belief'. Recall the feature of doxastic deliberation that I have been referring to as "transparency": the question *whether to believe that p* immediately gives way to the question *whether p is true* within the first-personal deliberative context. In part this means that once one has discovered whether *p* is true, the question *whether to believe that p* is *closed*.[36] We can test the adequacy of any proposed account of the meaning of 'belief' by seeing whether, once we substitute the proposed analysis for the term 'belief', the question remains closed. We might call the utilization of this variant on Moore's test to knock down proposed definitions of 'belief' the *closed question argument*.[37]

The teleological account before us falls prey to the closed question argument, because as we have seen, it is not a closed question whether, in asking oneself *whether to believe that p*, where belief is conceptualized

Published by Duke University Press

as being weakly responsive to evidence, the issue is resolved solely by thinking about *whether p is true*. According to the teleological account, it may be that one is obliged to see the truth of the proposition as relevant to settling what to believe, on pain of not seeing the belief as responsive to evidence at all, but one can satisfy the weak dispositional condition of being minimally responsive to evidence for and against *p*'s truth without taking the truth or falsity of *p* to *settle* the issue and thereby *close* the question of *whether to believe that p*. One may in fact discover that *p* is true, on the teleologist's account, but still not settle *whether to believe that p*, because one may judge that in the circumstances other considerations are more important than *whether p is true*. But treating non-evidential considerations as normatively outweighing the truth (or of having any normative relevance) in doxastic deliberation would violate the condition of transparency, which entails that once decisive evidence of the truth as to *p* is discovered, the question of *whether to believe that p* is closed, no other considerations being allowed to alter the deliberative outcome. Teleological accounts thus succumb to the closed question argument by leaving open a deliberative question that is in fact shut.

Could the teleologist argue that because deliberation is a *self-conscious* activity, an agent must think of himself as aiming at truth if he is to treat his reasoning as deliberation about what to *believe*, and that unless he treats evidence as solely relevant in his reasoning, he cannot see himself as aiming at truth? No, because the teleologist has told us that aiming at the truth requires having only a *weak* disposition to regulate one's cognition by evidence. Therefore, all the teleologist can require of someone in order that he conceive of himself as aiming at truth is that he treat evidence as relevant to determining belief, not that he treat it as *solely* relevant. The phenomenology of deliberation, however, is that evidence is the only kind of consideration that can provide a reason for belief, because only evidence is relevant to answering the question "Is *p* true?" that one finds oneself directly faced with in deliberation about *whether to believe that p*.

The only way for the teleologist to avoid this result is by requiring that the disposition to be regulated by evidence that is constitutive of belief be so strong as to overpower the potential influence of any other factors in belief-formation. If this were so, the only way an agent could deliberate about what to believe is if he were influenced solely by evidence. If he were to take non-evidential factors as relevant to settling belief, then he wouldn't be able to treat his deliberation as realizing the

Published by Duke University Press

strong disposition constitutive of belief-formation. While this solution might appear to account for transparency,[38] it is impaled on the other horn of the dilemma. For now the teleologist has saddled himself with such a strong disposition that he cannot make room for the paradigm cases of pragmatically influenced belief. If the disposition constitutive of belief were so strong as to rule out the influence of any other factors, then any cognition that came to be influenced by non-evidential factors would cease to be a belief. The wishful thinking cuckholded husband who believes his wife to be faithful or the racist whose dislike of African-Americans leads him to fall prey to poor sociobiological or other "scientific" arguments wouldn't be irrational, but impossible!

Again, to repeat, the phenomenon of transparency does not imply that evidence has sole *causal* influence over belief, but rather implies that evidence has sole *normative* or *rational* authority over belief. The strong dispositionalist proposal requires us to accept far more than transparency demands, transgressing our stable, widespread, and commonplace judgments of the existence of irrational, pragmatically induced beliefs.

Let me now attempt to extract a lesson from the failure of teleological accounts. Go back to the analogy that provides the impetus for teleological accounts—the analogy between belief-formation and other aim-constituted activities such as housebuilding or chess playing. The teleologist interpreted the analogy as showing that belief is a state that is constitutively regulated by an intention to accept the truth, or some surrogate of such an intention. But just as one can produce a house without attempting to provide shelter, one can form a belief even if one isn't intending to accept the truth. One can unwittingly build a house by intending merely to follow a set of instructions, and so too one can unwittingly form a belief, as when one forms beliefs by way of the operation of unconscious processes. We need an account that explains why one can't conceive of one's activity as inquiry unless one is disposed to treat evidential considerations, and only evidential considerations, as reasons for belief, but without building this disposition into the metaphysics of belief.

Belief as a Normative Concept

Possibly inadvertently, Velleman himself alludes to an alternative account of the way that truth is the standard of correctness for belief:

Published by Duke University Press

> The concept of belief just is the concept of an attitude for which there is such a thing as correctness or incorrectness, consisting in truth or falsity. For a propositional attitude to be a belief just is, in part, for it to be capable of going right or wrong by being true or false.[39]

Right after this Velleman claims that the hypothesis that belief aims at truth is an attempt to account for this inescapable authority that truth has over belief. This suggests that he doesn't think that he has already given a possible explanation of this inescapability in the previous quotation. However, if as Velleman claims, it is just part of the concept of belief that beliefs are correct if and only if they are true, then we have our explanation, albeit a trivial one. Unless one has grasped that truth bears this normative relation to belief, one will not have grasped the meaning of 'belief'. Thus, because it is a conceptual matter that truth is a standard of correctness for belief, it is unnecessary to look for further facts to explain how truth is inescapably normative for belief. The inescapability comes from *conceptual necessity*.

Of course, taking this route still requires explaining what it is to think that the standard of correctness for belief is truth. Unless we understand what it is to take truth to be a standard of correctness for belief, we won't have understood the concept of belief. But it is only if one has an antecedent commitment to deriving normativity from teleology that one will think that the only way that truth can be the standard of correctness for belief is if truth provides the *telos* for belief. Instead, if one thinks that normative language can be understood without reference to the purpose or design of the entities to which norms apply, the move from standards of correctness to aims or designs will appear in need of motivation. Given the dilemma that teleological accounts face in trying to account for the way that truth is normative for belief, I suggest we look elsewhere for our understanding of the conceptual necessity that ties truth normatively to belief.

We can work our way into my proposal by first reminding ourselves of the phenomenon that needs explaining. Deliberation of any kind is framed by a question, whether it is what to do, what to believe, what to pretend, or whatever. This does not mean that an agent has to have the question at the forefront of his mind, explicitly posing the question to himself, as it were; but unless his thinking manifests some recognition that this is the question that he is striving to answer, his stream of thought would lack the direction or purpose required for it to be an instance of deliberation about what to do or believe rather than, for example, a stretch of directionless cogitation. The phenomenon of

Published by Duke University Press

transparency that needs explaining occurs within the context of deliberation that is structured by the question *whether to believe that p*. Transparency does not occur in nondeliberative contexts of belief-formation, nor does it occur (at least in a sense that isn't trivial) in deliberation whose sole question is *whether p is true*. So if we are to account for transparency, what we need to explain is why deliberation that is framed by the particular question of *whether to believe that p* is answered solely by considerations relevant to answering the question *whether p is true*. We don't also need to explain instances of nondeliberative belief-formation or deliberation that is framed by some other question, because there is no phenomenon of transparency that occurs in those cases. As we saw, one thing that goes wrong with the strong dispositional version of the teleological account of belief is that it requires us to think that transparency occurs where it doesn't, and what goes wrong with the weak dispositional version is that it fails to account for transparency where it does occur. We need an account that explains why deliberative belief-formation is regulated solely by a disposition to be moved by alethic considerations, but doesn't require that nondeliberative instances of belief-formation be also solely regulated by such a disposition.

My proposed avenue of explanation thus comes into view when we recognize that transparency occurs only in the context of *asking* oneself what to *believe*. As I pointed out, this does not mean that an agent has to explicitly ask himself this question; all that is required is that the question be in the background of his reasoning, guiding his deliberation. What I suggest is that by framing his deliberation as answering to the question *whether to believe that p*, a disposition to be moved by considerations that he regards as relevant to the truth of *p* and a disposition blocking considerations that he regards as irrelevant to the truth of *p* are activated. That is, part of possessing the concept of belief involves being disposed in this way when one applies the concept to frame one's reasoning. But, on this view, when one doesn't exercise the concept of belief, as in cases of nondeliberative belief-formation, there is no guarantee that one's cognitive activity will be regulated by these dispositions. And this is exactly as it should be, because as we saw, the phenomenon of transparency occurs only with deliberative belief, in which an agent exercises the concept of belief; it doesn't also occur when belief-formation fails to involve the exercise of the concept of belief.

Published by Duke University Press

Nor is my view committed to the claim that an agent counts as forming a belief only if, on reflection, he would take evidence, and only evidence, as reasons for maintaining or abandoning the attitude. Whether an agent takes reasons of evidence to be solely relevant to the maintenance or abandonment of an attitude depends upon whether he thinks of that attitude as a belief. If he views the attitude as a belief, then he will take evidence to be solely relevant, but if he thinks of the attitude as something he is assuming for the sake of a pretense, for example, then he won't take evidence to bear at all on whether to maintain or abandon the attitude. To claim that an agent does not believe that *p* unless he would be guided, upon reflection, solely by evidence for or against *p*'s truth in determining whether to maintain the attitude toward *p* would seem to imply that an agent cannot form beliefs unless he has the sophisticated conceptual wherewithal needed to possess the concept of belief, and that an agent's state of mind does not count as a belief unless he would accept that it is a belief. These are quite controversial claims, as they seem to both exclude non-ratiocinative creatures from having beliefs and grant agents legislative authority over whether or not their attitudes count as beliefs. Whatever the merits of this proposal, I need not commit myself to it in order to explain transparency. My point has been to emphasize the relevance of the trivial point that in order to deliberate about what to believe, one needs to possess the concept of belief, and my proposal is that it is the dispositions constitutive of possessing the concept of belief, and of seeking to answer a question framed with that concept, that are responsible for the fact that only truth-regarding considerations move an agent in such deliberation. Whether a creature needs to be able to reason about what to believe in order to have beliefs is a separate question, one whose answer determines whether or not transparency applies to all or only some believers, but does not bear on the explanation of transparency.

Up to this point, my proposal amounts to little more than re-describing the phenomenon of transparency in a way that brings out the point that transparency is expressive of a *conceptual* truth about belief. However, this is an important point that can easily be missed. An agent's grasp of this constitutive truth about belief shows up phenomenologically in the way that the truth of *p* appears to him as solely relevant to settling *whether to believe that p*. If this constitutive truth about belief were merely a metaphysical truth, then it would be possible for an agent to fail to appreciate it, and it would be possible for the truth about belief to fail to influence his deliberation. But transparency is the consciously

Published by Duke University Press

felt authority of truth for belief in any deliberation that aims to settle belief, so transparency can't be the conscious face of a merely metaphysical truth about belief. Transparency thus must express a conceptual truth about belief; a truth that an agent grasps merely in virtue of treating his deliberation as answering to the question of *whether to believe that p*. The next part of my proposal is that this conceptual truth about belief is a normative truth.

We have already seen that we ought not to characterize doxastic deliberators as being moved by the descriptive thought that only truth-relevant considerations are capable of causing beliefs. That would be to portray anyone who uses the concept of belief as involved in an error, since we know that belief isn't always caused by solely evidential factors. Moreover, most, if not all, users of the concept of belief know this fact; so to characterize the thought constitutive of having the concept of belief in this way is to convict most, if not all, of us of having inconsistent views about the causes of belief.

In any case, the descriptive thought that only evidential considerations can cause belief really can't explain the *immediacy* of the causal influence of evidential considerations in doxastic deliberation. To see this, let us assume that this descriptive account of the concept of belief is true and that I am deliberating about *whether to believe that p*. Thus, I know that my deliberation won't count as belief-formation, and the conclusion I draw won't count as belief, unless the deliberation is solely influenced by evidence for and against *p*. This means I know that if my object is to form a belief about *p*, focusing on anything other than evidence of *p*'s truth (or falsity) will be ineffective. Expressing this to myself, I might think, "I'd like to form a belief about *p*. But my reasoning won't count as belief-formation unless I focus solely on whether *p* is true; therefore in order to achieve my goal, I had better go about discerning whether *p* is true." The focus on truth here is of *instrumental* value in the achievement of the goal of belief; therefore the focus on truth is *mediated* by an inference that it would be conducive to my goal to focus on truth. However, as I pointed out in the introduction, in doxastic deliberation there is no inference that mediates the transition between asking oneself *whether to believe that p* and asking oneself *whether p is true*. Thus, an agent's having the descriptive thought associated with teleological accounts of belief cannot explain the immediate relevance for the agent of evidential considerations in doxastic deliberation.[40]

I suggest that an agent's acceptance of truth as the standard for correctness for belief functions *prescriptively* in deliberation about *whether*

Published by Duke University Press

to believe that p, directing an agent to believe *p* only if *p* is true. It is when an agent is guided by this directive that the dispositions solely to be moved by evidential considerations are activated, influencing the agent's reasoning.[41] That this prescription is accepted in virtue of an agent's conceiving his deliberation as answering to the question of *whether to believe that p*, rather than *whether to assume that p* or *whether to imagine that p*, implies that it is one of the conditions for possessing the concept of belief that one accept the prescription to believe that *p* only if *p* is true.

Here then is how this normative proposal about the concept of belief accounts for transparency. To say that it is a conceptual rather than merely a metaphysical matter that truth is the standard of correctness for belief is to say that a competent user of the concept of belief must accept the prescription to believe that *p* only if *p* is true for any activity that he conceives of as belief-formation. Because one accepts this prescription insofar as one is deliberating about *whether to believe that p*, determining *whether p is true* will be immediately imperative, to the exclusion of any other question, for anyone who entertains the deliberative question *whether to believe that p*. Again, notice that if my normative proposal were merely a claim about the metaphysics of belief, then it wouldn't be able to account for transparency. It is because an agent grasps the normativity of truth for belief in his deliberation about *whether to believe that p* that he treats only evidence for and against the truth of *p* as a doxastic reason. If it were a mere fact that belief is subject to the standard of correctness of truth, though, an agent could fail to apprehend it, and thus the normativity of truth for belief could fail to move him in his deliberations about *whether to believe that p*. But every agent who deliberates about *whether to believe that p* recognizes the normative authority of truth for belief. That truth is belief's standard of correctness expresses a conceptual, not merely a metaphysical, truth, something an agent grasps just in possessing the concept of belief. In order to elaborate this proposal further, I will connect it to my discussions of Gibbard and Rosen.

Gibbard, although heeding Moore's lesson by not attempting a reduction of normative concepts to descriptive concepts, mislocates the source of the normativity that ties truth and evidence to belief. Instead of locating the normativity in the concept of evidence or truth, we ought to keep it where it belongs, in the concept of belief. The truth of a proposition does not provide the standard of correctness for all cognition, certainly not for imagination or supposition. And evidence

Published by Duke University Press

for the truth of a proposition provides no reason to assume the proposition for the sake of argument or to entertain in imagination the proposition as part of a game of pretense. Although we can assess the contents of these other cognitive states as true or false—after all, the bizarre but intriguing axiom that you assume in order to trace its implications may in fact turn out to be true and the dancing bear that you imaginatively entertain walking down the street may actually be walking down the street due to the fact that, unbeknownst to you, the circus is in town—truth functions as a *standard of correctness* only for *belief* and only evidence of truth provides a *reason* for *belief*.

As we saw earlier, Gibbard's failure to locate the source of epistemic normativity in the right place, **belief**, leaves him unable to explain satisfactorily why only evidential considerations can be reasons for belief and why evidential considerations don't provide reasons for other cognitive states. The hypothesis that applying the concept of belief to one's thinking involves an acceptance of truth as that to which one's thinking must conform, however, can account for the unique normative standing of evidence for belief. I suggest that we think of the prescription that one accepts in framing one's deliberation about *whether to believe that p* in terms of an objective norm. The objective norm of belief prohibits believing that *p* unless *p* is true. But since we cannot directly check our representations against the truth, we need subjective norms by which to guide our doxastic activities. The role of the objective norm, whose acceptance is expressed in the phenomenon of transparency, is to provide a standard of success for subjective norms of good evidence that an agent can directly apply in his doxastic deliberation. It thus is not a brute fact that subjective norms for rational belief must be evidential in character. Rather, this constraint falls out of transparency, which is just an agent's recognition of the authority of the objective norm of truth.[42] Furthermore, because the claim that truth is the standard of correctness for belief is a conceptual truth about belief, but is not a conceptual truth about any of the other cognitive attitudes, anyone who has the concepts of belief and the concepts of the other cognitive attitudes must recognize the unique normative status that truth-regarding considerations have for belief. So, according to my proposal, anyone who understands what a belief is in a way that allows him to distinguish beliefs from other cognitive attitudes must recognize both that evidence provides reasons only for the cognitive attitude of belief and that only evidence provides reasons for belief.

Published by Duke University Press

Such a proposal involves denying Rosen's claim that no constitutive standards of correctness, in themselves, underwrite reasons for conforming to them. At least with respect to belief, the experience of transparency within first-personal deliberation about *whether to believe that p* indicates that belief's standard of correctness does determine what counts as a reason for belief from within the first-personal deliberative point of view. And from a deliberative point of view, making a judgment about what one has reason to believe is bound to have motivational force, because the function of deliberation, as opposed to mere reflection, is to come to a decision or belief on the basis of one's appreciation of reasons. An agent, if he is to see himself as *deliberating* about *whether to believe that p*, must think that he is capable of forming a belief on the basis of the reasons that he discerns. Belief's standard of correctness, truth, thus does not function as a *mere* fact that an agent can ignore or be unmoved by in her deliberations about what to believe.[43] But neither does the standard of correctness always inform an agent's belief-formation, as shown by cases of practically induced beliefs. According to my proposal, the standard of correctness of truth solely governs an agent's thinking because he conceives of his mental activity as answering the question *whether to believe that p*. But my proposal is silent as to the causes of belief in contexts of nondeliberative belief-formation in which an agent does not apply the concept of belief. Thus, unlike the strongly truth-aimed teleological account of belief, my proposal does not transgress the platitude that beliefs are sometimes caused by non-evidential factors. But unlike the weakly truth-aimed teleological account, my proposal is able to account for the fact that evidential considerations are solely influential in one special context of belief-formation, deliberation about w*hether to believe that p*.

One might worry, though, that my proposal says nothing about normal, nondeliberative but evidence-driven cases of belief-formation, such as when one believes the deliverances of one's eyes. These cases don't involve the application of the concept of belief, even in the background of one's thinking, so I can't adduce the influence of the truth-regulated disposition involved in applying the concept of belief to explain such cases. I thus am left with no account of what are probably the most frequent cases of truth-regulated belief-formation.

Rather than being a defect of my view, this seems to me to be a virtue. Remember that the goal of my account was to explain the phenomenon of transparency. Normal cases of perceptual belief don't involve transparency, so they don't involve phenomena that my account was

Published by Duke University Press

meant to explain. We saw that the teleological account of belief failed precisely because it was forced to treat all cases of belief, deliberative and nondeliberative, alike, thus failing to capture what is special about cases of deliberative belief. My proposal attempts to remedy this defect by explaining transparency in terms of a feature unique to the context in which transparency occurs. This strategy thereby is guaranteed to avoid running together cases in which transparency occurs and cases in which it does not occur. Though my proposal is silent about nondeliberative evidentially caused cases of belief, this doesn't mean that the proposal can't be filled out in such a way as to explain such cases. For example, the weak truth-aimed disposition that Velleman's account of belief claims is constitutive of belief seems to explain why evidence-driven processes such as perception normally induce belief. Insofar as such a weak truth-aimed disposition can explain these cases, we have grounds for combining the normative and teleological proposals into a hybrid normative-teleological account of belief. The teleological part tells us that a disposition to be moved by evidence is activated in all cases of belief-formation, and the normative part tells us that in addition a strong disposition to block the influence of non-evidential types of influence is activated in cases of belief-formation that are governed by an agent's application of the concept of belief. My point earlier was that the teleological aspect of belief cannot explain the feature that is unique to cases of deliberative belief-formation, transparency, and therefore cannot tell us the whole truth about belief.

To see more clearly why we might combine normative and teleological elements, let us revisit our cuckholded husband one last time. If the husband turns his mind to the question whether to believe that his wife is faithful, then the concept of belief engages his thought, directing him to accept a proposition about his wife's fidelity only if he can discern its truth. If he were to accept some other norm as having jurisdiction over this episode of thought, then he could not consider his activity as one of belief-formation. This is part of what is meant by the thought that beliefs are states of mind that are correct if and *only* if they are true. If he never bothers to ask himself this question, however, then he may very well be induced by wishful thinking or other non-evidentially sensitive processes to be in a state of mind that, third-personally, we would judge to be the belief that his wife is faithful. Even the husband himself could come to judge that these practically induced cognitions are beliefs by taking a third-personal perspective on these cognitions, for example, in a therapeutic context. He would still be

Published by Duke University Press

required to think that his belief was defective, but so long as his cognition were shielded from the psychological force of his acceptance of the norm of truth, his cognition might well stay in place, unmoved by his assessment of reasons.[44] Of course, if his state of mind were completely immune to the causal influence of evidence—if, for example, even after repeatedly seeing his spouse in passionate embraces and other intimate encounters with another person, the husband had absolutely no tendency to change his mind—it probably would be odd to continue to hold that he believes, rather than pretends, that his wife is faithful. And the appropriate explanation for our unwillingness to call his propositional attitude toward his spouse's fidelity one of belief might be that it is constitutive of belief that it at least be weakly responsive to evidence.

Thus, we might have reason to accept both (1) that it is a conceptual truth about belief that truth is belief's standard of correctness, and (2) that it is a conceptual truth about belief that belief is minimally responsive to evidence. Because my goal in this paper has been to explain transparency, I have emphasized the merits of the first, normative, claim. Although I have argued that the second, descriptive, claim cannot be utilized in explaining transparency, for all I have said there may be other considerations in favor of accepting it as well.[45]

Conclusion

The argument for my normative proposal about belief came in two steps. First, I argued that it is in virtue of applying the concept of belief to frame one's deliberation as answering to the question of *whether to believe that p* that the strong dispositions to be moved only by considerations taken to be relevant to the truth of *p* are activated. Second, I claimed that the thought that activates these dispositions—a thought that is constitutive of possessing the concept of belief—is characterized in normative terms. Truth-regarding considerations are not inescapable for belief in the way that the laws of physics are inescapable for one's bodily movements. One doesn't buy into the laws of physics just by possessing the concept of action, but one does accept the sole authority of truth over one's cognition when one views it as a belief. This is just the point that transparency is expressive of a conceptual truth about belief. Another way in which the inescapability of the laws of physics for action differs from the inescapability of truth for belief is that the laws of physics determine the boundaries of what it is possible

Published by Duke University Press

to do, which our actions are incapable of violating, whereas truth provides an inescapable normative standard for belief that beliefs may or may not succeed in meeting. In doxastic deliberation an agent is not blindly influenced by evidence as a brute cause of belief, but rather takes in evidence as a reason for belief. The inescapability of truth for belief thus is both conceptual and normative.

At least one very large question that remains, even if my argument has been successful, is how to interpret the normative "correctness" thought that I have claimed is contained in the concept of belief. Various meta-normative views come to mind. For instance, conceiving of a mental state as subject to a standard of correctness can be understood as an attribution of a non-natural property to that mental state. But such an interpretation faces difficult though familiar metaphysical and epistemic questions (for example, What is the relation between this non-natural property and natural properties? And how do we gain epistemic access to this non-natural property?). Or this thought can be given a naturalistic reduction, although I have given grounds for doubting that a reduction of correctness in terms of dispositions to regulate states of mind by evidence can succeed. An intriguing and, to my mind, promising alternative route to understanding the normative force of treating truth as the constitutive standard of correctness for belief runs via norm-expressivism. Working out the details of such an analysis is a large undertaking involving well-known obstacles facing noncognitivist views in general, as well as more specific obstacles concerning whether a norm-expressivist analysis of the concept of belief is self-undermining or incoherent in some way. My purpose in this essay, though, has not been to adjudicate between different meta-normative interpretations,[46] but to suggest that we need to think of the concept of belief in normative terms if we are to explain successfully why the phenomenon of transparency holds within, and only within, doxastic deliberation. I leave the task of settling the question of how to interpret the hypothesis for another occasion.

Amherst College

Notes

I am grateful first and foremost to David Velleman for encouragment, discussion, and extensive written comments. I also would like to thank the following people for discussion of material presented in this paper: James Bell, Paul Boghossian, Philip Clark, Stephen Darwall, Alexander George, Allan Gibbard, Peter Gibbard, Danny Goldstick, Jay Leahy, Robert Mabrito, Casey Perin, Ser-

Published by Duke University Press

gio Tenenbaum, Kevin Toh, Maura Tumulty, and two anonymous readers for the *Philosophical Review*. I also am indebted to an audience at New York University for a stimulating discussion of this paper. Work on this paper was supported by a faculty research grant from Amherst College.

[1] See Richard Moran, "Making Up Your Mind: Self-Interpretation and Self-Constitution," *Ratio* n.s. 1 (1988): 135–51, at 146, and his book *Authority and Estrangement* (Princeton: Princeton University Press, 2001), 60–65. In the book (61), he argues that it is a mistake to think of "transparency" as expressing the claim that, for example, the questions, "Do I believe that there will be a third world war?" and "Is it true that there will be a third world war?" are indistinguishable from or reducible to each other, because the fact that answers the question about the war is different from the fact about a particular person's belief, and anyone asking himself these questions knows this. But he does not consider whether the deliberative question whether to believe there will be a third world war is indistinguishable from or reducible to the question of whether it is true that there will be a third world war, and whether it is the relations between these two questions that is meant by "transparency." It is quite plausible that, for this pair of questions, the fact that answers one question also answers the other question, since the deliberative question *whether to believe that p* is not answered by any psychological fact about oneself, and anyone who asks himself these questions must know that the questions are answered in the same way. If "reducibility" just means that the same facts answer both questions, and "indistinguishable" just means that an agent cannot think that the questions have different answers, then it may very well be that "transparency" refers to relations of reduction and indistinctness.

[2] It is natural to associate a use of 'ought' with the prescription to believe *p* only if *p* is true. But 'ought' has multiple uses, including an all-things-considered usage that may factor in considerations of desirability. Therefore, in order to avoid misunderstanding, from now on I will refrain from couching my proposal in terms of a directive ought, which might be confused with this other type of ought, and stick to speaking in terms of 'correctness', where accepting truth as the standard of correctness for belief involves accepting a directive to believe that *p* only if *p* is true.

[3] I will use the boldface **belief** to refer to the concept of belief, 'belief' to refer to the word, and belief (plain text) to refer to the property of being a belief, or when it is left open whether the metaphysics or the semantics of belief is being referred to. I may also omit the boldface when I use the phrase 'concept of belief'.

[4] Of course, whether we agree on which norm I am to accept will depend upon whether we share the same higher-order norms. That is, we must agree on the kinds of reasons that justify the acceptance of norms governing the fulfillment of professional obligations, and this is a matter of whether we accept the same higher-order norms for selecting lower-order norms.

[5] See Allan Gibbard, *Wise Choices, Apt Feelings* (Cambridge: Harvard University Press, 1990), chap. 3.

[6] Ibid., 46.

[7] It may be that, as ascribers of rationality to beliefs, we are also committed

Published by Duke University Press

to the beliefs we pick out as rational having some property that irrational beliefs lack. This may be due to our acceptance of a rule governing our practice of ascribing rationality that specifies that it is not allowable to ascribe a difference in rationality between beliefs without countenancing some difference in their properties, a difference capable of being picked out using a descriptive concept (so long as we have a name in our language for this property). See Gibbard, "Normative and Recognitional Concepts," *Philosophy and Phenomenological Research* (2002): 151–67, at 158–60, for an argument that the use of normative planning concepts, such as being *the thing to do*, commits one to thinking that there is a natural property that *constitutes* being the thing to do.

[8] Gibbard, *Wise Choices*, 36–37.

[9] Ibid., 37.

[10] Ibid., 222.

[11] Even the pragmatist William James appears to bow to the normative dominance of truth over belief in his assertion that there are only two duties with respect to belief, to search for truth and to avoid error. However, even though both duties cited are notably alethic in character (the imperative to avoid error itself expressing a fear of falsity or *untruth*), James's account of the nature of duties doesn't give him a principled way of restricting duties in belief in this way. If duties with respect to belief are merely expressions of passions, as James claims, then there is no reason that the passions that they express couldn't be purely prudential or moral in nature. Therefore, according to James's view, there is no reason that there couldn't be purely practical or moral duties with respect to belief. See William James, "The Will to Believe," in *"The Will to Believe," and Other Essays in Popular Philosophy* (New York: Dover, 1956), sect. 7.

[12] Of course, one might think utility is a guide to truth, but then an acceptance of this evidential norm would mediate the connection between judgments of utility and belief.

[13] That we accept that doxastic deliberation has this alethic normative character can be seen by the descriptions we think appropriate to give of examples like that of the cuckolded husband. If the husband claims that his belief that his wife is faithful is warranted because believing in her fidelity makes him feel good about his marriage, either he is making a bizarre claim about the evidential relevance of his feelings to the *truth* of the belief, or he is simply making a conceptual error about the concept of warranted belief. This is why we are apt to respond by pointing out that, as Gibbard would put it, he has succeeded in giving a reason for *wanting* to believe in his wife's fidelity; but alas, he is mistaken in thinking that he has even given a candidate *reason* for believing that his wife is faithful. Rather than trying to persuade him that he should adopt our norms for belief, charity advises us to attribute to him an acceptance of such pragmatic norms for a state of mind other than belief.

[14] Again, he could suggest that this would be a disadvantageous normative position, but he can't rule it out as a coherent conceptual possibility.

[15] Or we might re-interpret what he means by "belief" as something other than **belief**. This strategy would be forced upon us if we were to accept a Davidsonian constraint on interpretation, since according to Davidson it is a consti-

Published by Duke University Press

tutive feature of an interpretable agent that he understands the concept of belief and what constitutes a reason for belief. See Donald Davidson, *Inquiries into Truth and Interpretation* (Oxford: Oxford University Press, 1984).

16 Might the husband turn to transparency itself for support, arguing that pragmatic considerations actually do bear directly on truth? After all, his happiness augurs in favor of believing that his wife is faithful. Since transparency tells us that determining what to believe amounts to determining what is true, doesn't it follow that his happiness is evidence that his wife really is faithful? This line of thought suggests that the husband misunderstands the relation between truth and reasons for belief. It is not that we get our handle on the concept of truth by asking ourselves what considerations warrant belief, but rather that we get our handle on the concept of a reason for belief by understanding how the concept of truth sets an independent standard that it is the function of reasons for belief to track. If truth were merely *the good in the way of belief*, as some pragmatists seemed to think, there would be no independent way of fixing what reasons to believe there are, since whatever one judged it good to believe would count as evidence for the truth of the belief. But this is not the relation that truth bears to reasons to believe, since we do distinguish judgments of what it would be good or beneficial to believe from judgments of evidence for the truth of a belief. The fact that there is this distinction, combined with transparency, indicates that truth serves to anchor rational belief to evidence in a nontrivial way.

17 Gibbard, *Wise Choices*, 6–9. Gibbard uses the terms 'warrant' and 'rationality' interchangeably as names for his all-purpose normative concept. For the rest of this section I will stick with using the term 'warrant' for Gibbard's concept and use 'rationality' and 'reasons' for our commonsense concepts of rationality and reasons. As we shall see, I don't think Gibbard's concept of warrant coincides with our concept of rationality.

18 Fundamental, irresolvable, disagreement, of course, is an apparent feature of moral discourse that drew many to noncognitivism in the first place.

19 Gibbard, *Wise Choices*, 7. For example, if the content of the concept of rationality associated with morality permitted only actions that maximize happiness to count as morally right, then it would be conceptually impossible for there to be a moral disagreement about the correctness of utilitarianism. Of course, the possibility of substantive disagreement is compatible with accepting some minimal substantive constraints. We might think of these minimal constraints as ground rules that one must abide by if one is to be taken as a serious participant in the discourse in question, one whom it even makes sense to disagree or agree with. But unless the constraints contained in a concept still allow for normative disagreement, the discourse framed in terms of the concept will not have the characteristic that Gibbard thinks makes it apt for a norm-expressivist treatment.

20 The need for the normative concept of a reason arises only once there exist self-conscious agents who can ask themselves whether to act on their desires or believe their perceptions, the essence of reason judgments being to govern by way of answering first-personal deliberative questions. See Christine Korsgaard, *The Sources of Normativity* (Cambridge: Cambridge University Press,

Published by Duke University Press

1996) 92–94. Reason-judgments dispose one to believe because they engage one's deliberations about *whether to believe that p*. Any other influence on belief that they might have is hidden from the agent's first-personal point of view, failing to register in the agent's own understanding of the way reasons govern his beliefs.

[21] For a similar criticism of Gibbard, see Justin D'Arms and Daniel Jacobson, "Sentiment and Value," *Ethics* (2000): 722–48, at 743.

[22] Of course, if the concept of evidence is identical to the concept of a reason to believe, then such a person may be showing a lack of understanding of the concept of evidence. But this would still involve a lack of understanding of the concept of belief, since one cannot understand the concept of belief without understanding the concept of a reason to believe. Under the assumption that the concept of evidence is synonymous with the concept of a reason for belief, this would amount to requiring that a full understanding of **belief** require an understanding of **evidence**.

[23] See Gideon Rosen, "Meaning, Normativity, and All That" (unpublished manuscript, 1999).

[24] Ibid., 18.

[25] Ibid., 19.

[26] See David Velleman, *The Possibility of Practical Reason* (Oxford: Oxford University Press, 2000), 18–19.

[27] This intention must be a *de dicto* intention to build something under the mode of presentation provided by the concept **house**, not a mere *de re* intention to produce *this*, where 'this' refers to the model of a house provided by an architect's blueprint. Obviously, one can produce a house without intending that it be a house, as when one builds a house by following instructions specified in an unlabeled instruction manual, where one's intention in following the instructions is merely to discover what the manual is about. Such an activity may be correctly called house building, since it not an accident that one's intentional activity produced a house, but it will not be the case that one intentionally engaged in the activity of house building. This would require that one built the house under the control of a *de dicto* intention to build something conceived as a house.

[28] See Velleman, "On the Aim of Belief," in *The Possibility of Practical Reason*, chap. 11, 244–81, at 254: "An acceptance has the aim of being the acceptance of a truth when it is regulated, either by the subject's intentions, or by some other mechanisms in ways designed to ensure that it is true."

[29] For some discussion of the various ways that the aim of truth might be realized, see Velleman, "On the Aim of Belief," 252–53.

[30] There is an ambiguity in the term 'regulation'. When we say that a process is regulated by some mechanism, we might mean that the process is actually causally influenced by those mechanisms that are said to regulate it, or we might mean that those mechanisms could, in some sense of 'could', causally influence the process. If the latter is what is meant, then I am especially sympathetic, since I think that the concept of belief is tied to our practices of holding each other epistemically responsible, and that it is only appropriate to hold somebody epistemically responsible if they are capable of regulating

Published by Duke University Press

their cognitions by epistemic considerations. While we might apply the concept of belief even when such a capacity is absent, these will be derivative or metaphorical uses of 'belief'.

[31] As I shall argue later, what makes a cognitive state a belief is the following: one who is in that state is subject to the *directive* to be in it only if its content is true. Therefore, it may make sense to group together cognitive states that share certain motivational and/or inferential roles—rather than states that share the property of being regulated by truth-tracking mechanisms—under the category of belief, since the motivational and inferential roles are what determine whether it is *important* that a cognitive state be regulated by truth-tracking mechanisms.

[32] A teleologist might even allow for beliefs that fail to be regulated by any truth-tracking mechanisms, so long as some of the mechanisms that do regulate them were *designed* to ensure that they are true. For example, a mechanism might at one time have been truth-tracking, but due to some sort of malfunction, it now no longer is. If so, the teleologist could allow that the representational states that such a mechanism regulates still count as beliefs, because the *function* of the mechanism is to ensure that the representational states are accurate.

[33] Velleman is explicit about tailoring his account of the truth-regulated disposition constitutive of belief to allow for cases of non-evidentially influenced beliefs. See "On the Aim of Belief," 254.

[34] I leave it open what would count as evidencing such a susceptibility to evidence. It could be that merely attempting to refute the evidence would demonstrate such receptiveness, so long as the husband's refutation itself had some influence over his beliefs, and therefore wasn't merely rationalization.

[35] See G. E. Moore, *Principia Ethica,* rev. ed. (Cambridge: Cambridge University Press, 1996), 67–68.

[36] Transparency has the further implication that once one has determined *whether to believe that p*, it will be a closed, or at least settled, question for one *whether p is true*.

[37] Thanks to Steve Darwall for suggesting this apt label.

[38] I argue later on, though, that this appearance is deceiving, and that even the strong dispositional version of the teleological account cannot capture transparency.

[39] *The Possibility of Practical Reason*, 16.

[40] This constitutes my explanation of why even the strong dispositional version of the teleological account cannot account for transparency.

[41] What precisely is the relation between the prescriptive thought that one ought to accept only what is true and the truth-sensitive dispositions that are thereby activated? I think that how one answers this question will depend upon which meta-normative interpretation of the prescriptive thought that one favors. For example, if one favors a normative realist interpretation, then one might be apt to think that it is the perception of the normative fact that one ought to accept only what is true that causes one to be moved by solely truth-sensitive considerations. According to the norm-expressivist interpretation I favor, to think that truth is the standard of correctness of belief is to

Published by Duke University Press

accept a higher-order norm that rules out accepting non-truth-sensitive norms as reasons for belief. Thus, according to the norm-expressivist, the higher-order norm, which is expressed by the directive to believe only what is true, activates the truth-sensitive dispositions that embody one's acceptance of lower-order evidential norms given expression in one's assessments of reasons for belief. I plan to explore this kind of noncognitivist interpretation of the normative aspect of belief on a subsequent occasion.

[42] Of course, it is contentious among epistemologists which norms of evidence to adopt, but it is not contentious that the point of such norms is the attainment of truth. Therefore the specific determination of subjective norms of evidence by the objective norm of truth needn't be thought of as solely a conceptual, logical matter, because it may very well be a substantive, a posteriori matter as to which methods of inquiry are most likely to arrive at the truth.

[43] Maybe this just means that I am not talking about *correctness* in Rosen's sense. But it does seem pretty clear that there is a sense of correctness in play here that has a normative, reason-generating role in doxastic deliberation. In order to avoid confusion, it may be better to speak of the standard of truth as an *objective ought* that generates norms of good evidence that themselves furnish agent-accessible *subjective oughts*, leaving out talk of correctness altogether. But as I mentioned earlier (note 2), these *directive* oughts might themselves be confused with other uses of 'ought'.

[44] Our (and his therapeutic) grounds for thinking that the husband believes his wife to be faithful would be the usual kind, such as the role that such a state has in the guidance of his behavior.

[45] In fact, we might think that the descriptive, teleological claim falls out of the normative claim. After all, in general it seems pointless to subject someone to a norm if they cannot even attempt to comply with it. This suggests the possibility of a unified way of combining these two elements into a normative-teleological account of belief, the normative element giving the rationale for the teleological element.

Furthermore, there is a way in which teleology might explain the normative element in belief. In giving an account of the normative element in the concept of belief, we will want an explanation of why we apply evidential norms (and only evidential norms) to certain cognitive states. Those of us who are inclined toward a nonrealist account of normativity, unwilling to cite a capacity for appreciating normative facts as an explanation of our normative thoughts, might look for an evolutionary rationale of the adaptive utility of governing certain types of cognitive states by truth-sensitive norms. If such an evolutionary explanation could be achieved, in one sense we would have accounted for the normative by way of the teleological. I plan to explore these possibilities for combining teleological and normative elements in an account of belief on a subsequent occasion.

[46] Nor is my purpose to rule out a line of thought that questions the presupposition that there is a need for a philosophical interpretation of normative commitments. A certain type of Wittgensteinian, for example, might think the analysis that I have given is fine as it is, and that it is only an illusion induced by philosophy that makes the analysis appear in need of a legitimizing interpreta-

Published by Duke University Press

tion. As far as anything I have said in this essay about my analysis of belief, quietism might be the correct attitude: there may well be no need for further philosophical work to make the analysis fully intelligible.

Published by Duke University Press