

## Explaining doxastic transparency

Sullivan-Bissett, Ema

DOI:

[10.1007/s11229-017-1377-0](https://doi.org/10.1007/s11229-017-1377-0)

License:

Creative Commons: Attribution (CC BY)

*Document Version*

Publisher's PDF, also known as Version of record

*Citation for published version (Harvard):*

Sullivan-Bissett, E 2018, 'Explaining doxastic transparency: aim, norm, or function?', *Synthese*, vol. 195, no. 8, pp. 3453–3476. <https://doi.org/10.1007/s11229-017-1377-0>

[Link to publication on Research at Birmingham portal](#)

### General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

### Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact [UBIRA@lists.bham.ac.uk](mailto:UBIRA@lists.bham.ac.uk) providing details and we will remove access to the work immediately and investigate.

# Explaining doxastic transparency: aim, norm, or function?

Ema Sullivan-Bissett<sup>1</sup>

Received: 5 July 2016 / Accepted: 12 March 2017  
© The Author(s) 2017. This article is an open access publication

**Abstract** I argue that explanations of doxastic transparency which go via an appeal to an aim or norm of belief are problematic. I offer a new explanation which appeals to a biological function of our mechanisms for belief production. I begin by characterizing the phenomenon, and then move to the teleological and normative accounts of belief, advertised by their proponents as able to give an explanation of it. I argue that, at the very least, both accounts face serious difficulties in this endeavour. These difficulties are a function of seeking an explanation of transparency at the agential level, either with the subject aiming at truth, or being guided by a norm of truth. I adopt a motivational account of belief, one which severs the connection between belief and truth, and supplement this with an account of actual world beliefs. My alternative explanation is found at the sub-intentional, non-agential level, secured by biology. This explanation casts transparency not as related to the nature of deliberation over what to believe, but rather as *contingently* characterizing the beliefs of *some believers*, namely those with a particular biological history. My explanation thus parts company with what has come before along two dimensions: it moves away from transparency being something related to the agent's aims or commitments, and it understands it as a contingent phenomenon. I close by considering an objection to my view—that transparency must not be understood as a contingent phenomenon—and a nearby alternative position which avoids this consequence. I respond to this objection and give reasons not to endorse the nearby alternative. I conclude that my explanation does not face the difficulties of those offered by teleologists and normativists, and, that by moving away from agential explanations, and casting transparency as contingent, we can provide a successful explanation of it.

---

✉ Ema Sullivan-Bissett  
e.l.sullivan-bissett@bham.ac.uk

<sup>1</sup> Department of Philosophy European Research Institute, University of Birmingham, Edgbaston, Birmingham, West Midlands B15 2TT, UK

**Keywords** Transparency · Aim · Belief · Norm · Function · Truth

## 1 Transparency

Here I am interested in the best explanation of *transparency*.<sup>1</sup> I will not argue that transparency holds, but assume that it does, and focus on putative explanations of it.<sup>2</sup> Transparency characterises doxastic deliberation (that is, deliberation structured by the question *whether to believe that p*): ‘when asking oneself *whether to believe that p*’ one must ‘immediately recognize that this question is settled by, and only by, answering the question *whether p is true*’ (Shah 2003: p. 447). This is not the relatively weak claim that the truth of *p*, by the subject’s lights, is required for her to believe that *p*, though does not suffice as grounds for believing that *p*. This would leave it open that there might be non-truth related grounds, perhaps to do with the practical consequences of believing that *p*. Transparency does not leave this open. Rather, the question *whether to believe that p* collapses into the question *whether p is true*. Truth (by the subject’s lights) is not only *necessary* for belief when deliberating over what to believe, it is also *sufficient*. If a subject resolves *whether p is true* affirmatively, then she believes that *p*.

The step from the deliberative question *whether to believe that p* to the question *whether p is true* is ‘immediate and not inferential’ (Engel 2007: p. 198). When we deliberate over *whether to believe that p*, the truth of *p* is ‘not an optional end for

<sup>1</sup> Another kind of transparency has been identified and discussed by Richard Moran (1988, 2001). Transparency of this kind is one of self-knowledge: the question ‘Do I believe that *p*?’ is not distinguished from, and is answered by, answering the question ‘Is *p* true?’ (Moran 1988: p. 142). This is transparency about what I *do* in fact believe, and not transparency about what I *should* believe (in cases in which I have not yet formed a belief). I am interested in the latter kind of transparency here, and so I put the self-knowledge version aside. (For discussion contrasting the two, see Shah 2003: p. 476, n.1.) Nishi Shah and J. David Velleman suggest that this latter kind of transparency is the fundamental variant, noting that ‘what is *fundamentally* transparent to the factual question *whether p* is the deliberative question *whether to believe that p*’ (Shah and Velleman 2005: p. 518, my emphasis). I take no stand on the issue of fundamentality, my aim is rather to offer an alternative explanation of transparency in light of the shortcomings of previous explanations.

<sup>2</sup> Many philosophers would grant this assumption, and accounts of the connection between belief and truth have been offered in part as a way of accommodating the truth of it. Recently though there has been some dissent, with some philosophers arguing that transparency does not characterize doxastic deliberation (e.g., McHugh 2012a, 2013b; McCormick 2015). Elsewhere I respond to these discussions, arguing that putative non-transparent cases of deliberative belief formation (in which it is said that non-evidential considerations enter as *reasons* for belief), can be accommodated by the transparency proponent. Such cases, I suggest, are ones in which non-evidential considerations merely change the standards required for believing in a particular context, they do *not* provide reasons for belief which are considered *as such* from the deliberative perspective (Sullivan-Bissett 2017a, forthcoming). For those not persuaded that transparency characterizes doxastic deliberation, I invite them to read the main claim in this paper as the following conditional: if transparency did characterize deliberative belief formation, then teleological and normative explanations face difficulties, but it could be explained by appeal to a biological function of our mechanisms for belief production. For those whose scepticism about transparency is driven by concerns regarding explanations given by aim and norm accounts, perhaps such scepticism can be alleviated by my alternative proposal.

first-personal doxastic deliberation, providing an *instrumental* or *extrinsic* reason that an agent may take or leave at will' (Shah 2003: p. 447). If it were then the agent would have to make an inference from discovering that *p* is true and determining whether she ought to believe it, perhaps via what Shah calls a 'bridge premise' relating to whether or not it is good to have a true belief with respect to *p*. But there is no such inferential step in doxastic deliberation. The two questions then—*whether to believe that p* and *whether p is true*—are inseparable; one cannot take oneself to have answered the latter, without also taking oneself to have answered the former.

It is not required that the deliberating subject explicitly pose to herself the question *whether to believe that p* for her deliberation to exhibit transparency. The requirement is only that the subject's thinking 'manifests some recognition that this is the question' that she is seeking to answer, for without this manifestation, her stream of thought would not be deliberation over what to believe, but would be better described as a 'stretch of directionless cogitation' (Shah 2003: p. 466).<sup>3</sup>

So far, the characterization of the phenomenon is one I fully accept. Its supposed modal status however, is something which will not be accommodated by my explanation. The modal status of transparency has been taken to be strong, it is not merely 'a quirky feature of human psychology', a contingent feature of human minds (Shah 2003: p. 447). Rather, it has been taken to be something in the very nature of deliberation over what to believe which demands transparency. We will see later that on my view transparency is a contingent feature of actual world deliberative belief formation, and I will argue that this is no cost (Sect. 5.3).

What we are after then, is an explanation of why the question *whether to believe that p*, collapses into the question *whether p is true* in doxastic deliberation. In the next two sections I consider two broad kinds of explanation currently on offer in the literature, and argue that they are both problematic. The discussion here will be relatively brief, since the pay-off of the paper will be tendered by my alternative proposal. Given this, I may not do enough here to persuade the reader that my objections to the explanations of transparency offered by the teleological and normative accounts are insurmountable (though I think that they are, and I would gladly accept agreement on that point). I hope though to shore up enough scepticism about these explanations that an alternative position is welcome. Some of what is said in the next two sub-sections is not new (I refer the reader to more comprehensive formulations of these objections in other work). However, what has not yet been noted is what these objections reveal, namely, that it is problematic to explain transparency at the agential level, with the believer playing a role in the structuring of doxastic deliberation.

<sup>3</sup> Conor McHugh has suggested that the claim of transparency is surely too strong, if it is intended as ruling out that a thinker 'cannot so much as be struck by pragmatic considerations' (McHugh 2013a: p. 449). Rather, McHugh says, we should make the claim of *exclusivity*, which says that even if pragmatic considerations strike the deliberator, they cannot be motivationally efficacious. I agree with McHugh on this point, though I will put it aside and continue to talk in terms of explaining *transparency*, as long as we do not think of this as excluding practical considerations from *entering consciousness*, but only as excluding such considerations from *having efficacy from the subject's perspective* if they do so.

## 2 The teleological account

The teleological account of belief holds that belief is constitutively aim-governed. My interest here is in giving an explanation of the transparency to truth considerations in doxastic deliberation, and so I will focus on the *truth*-aim teleological account,<sup>4</sup> which has been presented as able to do so.

First, a note on aims. Normativists can employ aim-talk metaphorically—I postpone discussion of such accounts until section three. Here I am interested in aim accounts which are not metaphorical, ones which import our ordinary notion of aims, and explicitly eschew normative talk in favour of descriptive claims about the nature of belief.<sup>5</sup> The difference between teleological and normative accounts is nicely illustrated with an example from Asbjørn Steglich-Petersen: John says ‘I believe that the London train leaves at 5 p.m.’. In addition to judging that John regards as true the proposition *the London train leaves at 5 p.m.*, teleological and normative accounts claim that we have made an additional judgment. The former has it that we judge of John that he has an *aim* of regarding this proposition as true only if it is true. The latter has it that we judge of John that he *ought* to regard this proposition as true only if it is true (Steglich-Petersen 2006: p. 499).

Velleman (2000) and Steglich-Petersen (2006, 2009, 2011) have defended this account, according to which ‘believing that *p* essentially involves having as an aim to believe *p* truly’ (Steglich-Petersen 2009: p. 395, my emphasis). This aim is realized in one of two ways. In the deliberative case it is realized in the subject’s personal level intentions, and in the non-deliberative case it is realized by ‘some sub-intentional surrogate of such intentions in the form of truth-regulated [...] mechanisms’ (Steglich-Petersen 2006: p. 510). It is taken as both a necessary and sufficient<sup>6</sup> condition on an attitude being a belief, that it is aimed at truth in one of these ways.

### 2.1 The teleological account and transparency

Steglich-Petersen claims that we can explain transparency by appeal to the truth-aim of belief. When we deliberate over *whether to believe that p*, transparency ‘can be explained by the aim one necessarily adopts in posing that question, because the only considerations that could decide whether believing *p* would further that aim are

<sup>4</sup> At least one other aim has been put forward by McHugh, the aim of knowledge (2011). I do not discuss this account here since, as I noted above (fn. 2) McHugh denies that transparency places an explanatory burden on the teleologist, and so does not take his account to be explanatory of it. Importantly though, if one did think the aim of knowledge could explain transparency, the worries I consider in this section are concerned with beliefs being aimed at all, not how the aim is spelled out, and so may well provide traction against a knowledge aim account too.

<sup>5</sup> Ralph Wedgwood notes that the claim that beliefs aim at the truth must be taken to be metaphorical since ‘[b]eliefs are not little archers armed with little bows and arrows: they do not literally “aim” at anything’ (Wedgwood 2002: p. 267). Although the teleological account does not commit to this bow-and-arrow conception of belief’s aim, it does import our ordinary notions of aims, and means not to use this term merely metaphorically.

<sup>6</sup> David Owens (2003) has argued, correctly in my view, that guesses aim at truth. If this is right, the teleologist should let go of her claim that aimedness is sufficient for belief and/or distinctive of it.

considerations that bear on whether  $p$  is true' (Steglich-Petersen 2008: p. 546). This is an explanation of a modally strong reading of transparency because belief, it is argued, is *essentially* governed by the aim of truth. What is said to explain the move from the question *whether to believe that  $p$* , to the question *whether  $p$  is true*, is to do with the essential nature of belief—its having the aim of truth. If belief necessarily aims at truth, and this aim structures doxastic deliberation such that it exhibits transparency, then transparency is necessary to deliberation over what to believe.

That this explanation of transparency faces difficulties can be shown with the following dilemma: on one horn, the putative aim of belief cannot be weighed against other aims and thus is not appropriately thought of as an aim at all. On the other horn, we grant that the aim of belief can be weighed, and so there is such an aim, but it cannot explain transparency. So either there is no aim of belief, in which case there can be no explanation of transparency by appeal to it, or there is an aim of belief, but it is not at the required strength to explain transparency. Either way, aim-talk is not explanatory.

Let us turn to the first horn, that the putative aim of belief is not an aim after all, as it is not something which can be weighed. Owens characterizes the truth-aim thus: 'Ø-ing that  $p$  aims at the truth if and only if someone who Øs that  $p$  does so with the purpose of Ø-ing that  $p$  only if  $p$  is true' (Owens 2003: p. 289). Owens notes that he has characterized the truth-aim negatively; that is, an agent is only required by it to avoid error, she need not seek to believe all or any particular truths (Owens 2003: p. 289). This is important: if the claim of the teleological account was that we aim to believe that  $p$  if and only if  $p$  is true, we would find ourselves attempting to meet an impossible aim (there are simply too many truths—some of which are inaccessible to human believers).<sup>7</sup>

Owens argues that the aim of belief as given by the teleologist does not cohere with our ordinary notion of aims, since we understand aims as capable of interacting with and being weighable against each other. Here Owens places a constraint upon aims which is derived from the domain of action. The aim one might have with respect to some activity can interact with and be weighed against one's other aims. For example, my aim to run 5 km this evening is constrained by some of my other aims, including but not limited to, my aim to finish writing this section before bedtime. If I am a little behind on this section, I may decide to run just 3 km, if I am very behind on this section, I may discard my aim to run 5 km altogether and remain at my desk.

So if we want to appeal to some aim we have in believing, we ought to be able to weigh that aim against other aims we have with respect to the beliefs we form. Owens's claim is that this necessary condition on aiming is not met by the putative aim of belief. If I aimed to believe  $p$  only if  $p$  were true, then, given the nature of aims, I could weigh my aim to believe truly, against say, my aim to be happy. I might weigh my aim to have a true belief regarding my partner's fidelity against my aim to maintain emotional stability, and, if the latter aim is accorded more weight, I may end up not forming the belief that my partner is unfaithful (when my evidence suggests that he is). However, transparency blocks this kind of weighing. When we

<sup>7</sup> An analogue of this kind of worry arises for the formulation of the norm governing belief posited by normative accounts (Sect. 3).

deliberate over *whether to believe that p*, we are inescapably moved by our answer to the question *whether p is true*, regardless of any non-epistemic aims we might have when we enter this process. Owens claims that this exclusion of our other aims which occurs in deliberative belief formation shows that belief is not an aim governed activity<sup>8</sup> (Owens 2003: pp. 298–299). If there is no aim of belief, we cannot appeal to one in an explanation of transparency.<sup>9</sup>

A possible response at this point is to put pressure on the claim that aims are necessarily weighable. If one does not find this thought persuasive, one may well allow for an aim of belief *even though it cannot be weighed*. Here is not the space to give a full account of aims and the conditions thereon, I just make two quick points. First, in response to this argument, teleologists have gone one of two ways: claim that aims can be weighed in other deliberative contexts (Steglich-Petersen 2009), or claim that the aim can be weighed *in the context of* deliberation, since transparency does not characterise all such deliberation (McHugh 2012a).<sup>10</sup> Proponents of aim accounts then have not so far taken issue with Owens's characterisation of aims as necessarily weighable, and this might be suggestive of such a route being an unproductive one. Second, all other aims, can, at least in principle, be weighed. Even those which structure most of our activities and ones we are loathe to give up, can, in the right circumstances, be weighed against our other aims. For example, we might aim to *stay alive*, but even this can be weighed in certain circumstances (say, trading one's life for that of a loved one). So if the teleologist wanted to respond to Owens's problem by claiming that belief does have an aim *even though it cannot be weighed*, then she will have to concede that this is different from any other aim, along at least this dimension. Perhaps that is a tolerable cost, but the teleologist at the very least owes us an account of what makes it an *aim*, and why it is different from all other aims. I suspect that answering these questions may well put pressure on whether talk of aims can be anything but metaphorical, which was precisely Owens's concern.

I now move to the second horn of the dilemma which requires us to grant, for the sake of argument, that the aim of belief can be weighed. I will show that the weighing of the truth-aim for belief is incompatible with the presence of transparency.

If it were an aim of mine which explained why I moved from the question *whether to believe that p* to *whether p is true*, I ought to be able to resist this move in some circumstances, and weigh the aim of belief against other aims I might have which are relevant to *whether to believe that p*. I deliberate over whether to believe that *my partner is unfaithful*. I have a lot of evidence that he is unfaithful, and if I wish to achieve my

<sup>8</sup> We might only get something weaker from Owens's argument: that there is no *truth* aim of belief, but there is some other aim which is weighable. However, it is enough for my purposes to show that there is no aim which can be unproblematically appealed to in order to explain transparency.

<sup>9</sup> Steglich-Petersen has replied to this objection by appealing to other deliberative contexts in which the putative aim of belief can be weighed. I respond to this elsewhere, arguing that the response fails, and concluding that Owens's objection remains unanswered (Sullivan-Bissett and Noordhof 2013). Steglich-Petersen has since replied to this (2017), and I again to him (Sullivan-Bissett and Noordhof 2017). I thus take the first horn of the dilemma here to be well-supported.

<sup>10</sup> I am interested in the best explanation of transparency in this paper, and so have omitted discussion of McHugh's response to Owens which goes via a denial of transparency. I argue against this response elsewhere (Sullivan-Bissett 2017a), see fn. 2 for an overview.

aim to believe truly I will form the belief that *he is unfaithful*. However, I have another aim, that of maintaining emotional stability until I have finished writing this paper, and if I wish to achieve my aim to remain emotionally stable, I will not form the belief that *he is unfaithful*. Let us say that achieving my aim to remain emotionally stable for the sake of paper completion is far more a priority for me than forming a true belief on the matter of my partner's (in)fidelity. Then, if it is an aim which explains the character of belief formation, there is no reason why, in the context of doxastic deliberation, I could not forego the truth-aim in favour of an aim relating to my emotional stability. Now, the actual character of our deliberation, insofar as it is characterised by transparency, does not allow us to do this. If it did, our deliberative belief formation would not exhibit transparency after all: the question *whether to believe that p* would not collapse into, and be answered by, *whether p is true*. So if belief had a weighable aim, then other considerations could weigh into the deliberation which results in the fixation of belief. Transparency though is the immediate and non-inferential collapsing of one question into another, and so it cannot be an aim which explains this. So now we have the second horn of our dilemma: if there were an aim of belief, it would not be at the required strength to explain transparency.<sup>11</sup>

The mistake of the teleologist is to suppose that transparency results from personal-level aims, and that we, as believers, play an active role in its governing our doxastic deliberation. But in uncritically importing the notion of *aims*, the teleologist also imports a problematic necessary feature of them. From this we learn that the putative aim of belief is not an aim after all, or if it were, it would not generate transparency. Better to understand transparency, I say, not by appeal to agential aiming, but as a phenomenon to wish we passively respond.

Shah raises a related problem for the teleological account, the so-called *teleologist's dilemma* (2003: pp. 460–5). In brief, the problem is that a descriptive claim about the nature of belief (it being regulated for truth as a result of an intentional aim or sub-intentional aim), will either be too strong to accommodate non-deliberative belief, or too weak to explain transparency. If we think in terms of truth regulation, in order to count some beliefs (like those arising from self-deception) *as beliefs*, the teleologist might want to characterise this truth regulation weakly. But then, appeal to this weak truth-regulation will not be able to explain the primacy of truth in doxastic deliberation. On the other hand, if the teleologist sought to explain transparency by casting the truth regulation as very strong, then she would rule out some other cases of belief *as beliefs*.<sup>12</sup> Again, I take the teleological account's vulnerability here to be a function of its appeal to the agential level in the search for an explanation of transparency. In so doing it fails to capture the agent's relatively passive response to truth suggestive of biological function (as I shall argue), rather than settled by anything at which the agent can be said to be aimed.

<sup>11</sup> Indeed, it has been recognised that the aim theorist cannot both keep exclusivity to truth considerations and the ordinary notion of aims as things which are weighable. McHugh, in his response to this problem, rejects the first in favour of the second (McHugh 2012a, 2013b).

<sup>12</sup> Steglich-Petersen has responded to this problem by appealing to the descriptive characteristic of *weak truth regulation*, which he takes to be secured by the aim of truth (2006: pp. 509–16). Elsewhere I respond to this by putting pressure on the claim that all beliefs can be grouped together by appeal to their sharing some regulatory feature (Sullivan-Bissett and Bortolotti forthcoming).

To conclude this section: if belief has an aim, this ought to be weighable. If it is not weighable, there is no such aim, and so the teleologist has no resources with which to explain transparency. If it were weighable, this would allow for cases of failure to answer the question *whether to believe that p* by considering only *whether p is true*. If such cases were possible, then transparency would not hold. And so the aim could not explain transparency. In addition, if the teleological account allows itself the resources to explain transparency by explicating the descriptive requirement of belief in terms of truth regulation as very strong, then she cannot accommodate many attitudes *as beliefs*. If on the other hand the truth regulation is weak enough to accommodate such cases, it will not be at the required strength to explain transparency.

These objections cluster around an underlying problem. Namely, it is a mistake to try to explain transparency by appeal to something at which the agent can be said to be aimed.

### 3 The normative account

Normative accounts claim that belief is constitutively normative, and have taken belief's normative dimension—in particular its being governed by norms relating to its formation—to be key to explaining transparency. Whilst such accounts can be understood as endorsing the claim that *belief aims at truth*, they do so only in a metaphorical spirit. Despite considerable debate over how to characterize the norm governing belief which can do this explanatory work, there is broad agreement that the most plausible characterisation of the norm must be parsed in terms of permissibility to believe *p* if and only *p* is true, and not an obligation to do so.<sup>13</sup> However, once the normativist moves to this strength, she cannot explain transparency by appeal to the norm. I have given this argument at length elsewhere, as well as responses to the various moves normativists might make in response (Noordhof and Sullivan-Bissett m.s.), and so I rehearse the argument only briefly in the next section to motivate the need for my alternative explanation of transparency given later.

Bykvist and Hattiangadi (2007) claim that a norm of belief which *requires* believing that *p* if and only if *p* is true would be too demanding. It would place too much strain on our relatively limited cognitive faculties, since it would follow that one ought to believe every true proposition. As the number of true propositions available to be believed is infinite, and there are propositions too complex for us to believe (for example, huge conjunctions of true propositions), the requirement of a norm formulated this way cannot be met (Bykvist and Hattiangadi 2007: p. 279).

<sup>13</sup> Krister Bykvist and Anandi Hattiangadi—who argue that the truth norm had better be parsed in terms of permissibility—take there to be a residual problem, even once the normativist takes this on board. This is one which results from blindspot propositions, those propositions which are impossible to believe truly (e.g. *there are no believers*) (2007: 281ff; 2013: 107–21). McHugh, in responding to Bykvist and Hattiangadi, raises further problems with accounts which have the norm of belief as prescriptive, as opposed to evaluative (McHugh 2012b: pp. 12–15). I want to give normativism a fair hearing on whether it can explain transparency, and so I will put aside Bykvist and Hattiangadi's, and McHugh's arguments for its blanket untenability. If they are right (and I think that they are), I would have an easy route to the conclusion I want in this section. To be sure though, let us take the harder route.

A characterization of the norm which avoids this consequence is one which moves from obligation to permissibility:

It is permissible to (believe that  $p$ ) if and only if  $p$  is true. (See for example [Shah and Velleman 2005](#): p. 519; [Whiting 2010](#): p. 216)

This norm forbids a subject from believing that  $p$  when  $p$  is false, but it only *permits* a subject to believe that  $p$  when  $p$  is true, and so it is not too demanding in the sense identified by Bykvist and Hattiangadi. The norm of belief then must be parsed in terms of permissibility, lest it place on us the obligation to have every belief that would be true. Importantly for my purposes, the norm formulated in this way has been advertised as able to explain transparency ([Shah and Velleman 2005](#): p. 519), in the next section I argue that it faces difficulties in this endeavour.

### 3.1 The normative account and transparency

Moving to permissibility, as the normativist must, means that her explanation of transparency is problematic. Whilst transparency might be explainable by appeal to a norm if that norm is one which requires the believer to believe *all truths*, the move to permissibility may well sacrifice these explanatory gains.

The explanation of transparency offered by [Shah \(2003\)](#) and [Shah and Velleman \(2005\)](#) is the following: when an agent deliberates on the question *whether to believe that  $p$* , this deliberation is framed by the prescription to believe that  $p$  only if  $p$  is true. The agent expresses her commitment to this prescription in doxastic deliberation, which activates two dispositions: one to be moved by considerations which are, by her lights, relevant to the truth of  $p$ , and one which blocks considerations which are, by her lights, irrelevant to the truth of  $p$  ([Shah 2003](#): p. 467; [Shah and Velleman 2005](#): p. 519).

This explanation of transparency faces the following problem: if the norm governing belief formation were that I *ought* to (believe that  $p$ ) if and only if  $p$ , then, once I recognize that  $p$  in deliberation, I move immediately to believe that  $p$ . This norm might explain transparency. But we have seen that this cannot be the norm of belief. Rather, when a subject enters doxastic deliberation, she accepts that she is *permitted* to (believe that  $p$ ) if and only if  $p$  is true, and if it is this which she recognizes, then such recognition cannot explain transparency. It is consistent with the subject recognizing that she is permitted to (believe that  $p$ ) if and only if  $p$  is true, that she could answer the question *whether  $p$  is true* in the positive, and fail to believe that  $p$ . But this is not possible in doxastic deliberation.

Let us return to my earlier example. I deliberate on whether to believe that *my partner is unfaithful*. There is a lot of evidence which points to his infidelity, such that I judge that it is true that he is unfaithful. The norm tells me that I am permitted to believe that *my partner is unfaithful*. But, as before, I have a strong aversion to forming this belief (I really would like to finish this paper), and given that the norm is one which only *permits* me to form it, and does not require me to, I ought to be able to not form the belief that *my partner is unfaithful*. But this is impossible.

A norm governing belief parsed in terms of permissibility is set at the wrong strength to explain transparency. This is because if I answer the question *whether p is true* in the positive, I recognise that I *may* believe that *p*, not that I *ought* to believe that *p*. Why think that mere permissibility would always, and inescapably, move me from an answer to *whether p is true* in the positive, to believing that *p*? Failing to believe that *p* in such a case is consistent with my belief formation being governed by the norm that I am permitted to (believe that *p*) if and only if *p* is true. If beliefs are merely *permissible* when true, why would I allow deliberation about *whether to believe that p* to be entirely constrained by *whether p is true*?

Could the permissibility normativist claim that once in the game of deliberating over *whether to believe that p*, one is committed to forming a belief with respect to *p*, and so will immediately move from judging that *p* to belief that *p*? In some cases, perhaps. But even when I am committed to forming a belief with respect to *p*, and I answer the question *whether p is true* in the positive, this ought still to be compatible with my failing to believe that *p*, since I am merely permitted to do so. If I am committed to forming a belief on the matter, then I might, as a matter of fact, move from affirmatively answering *whether p is true* to believing that *p*. But if the norm governing my belief formation is merely permissive, I should be able to, as a matter of *psychological possibility*, answer *whether p is true* in the positive, and yet fail to believe that *p*, in cases where my commitment wanes, or I do not much like the answer yielded regarding the truth value of *p* and so drop my commitment to forming a belief on the matter.

The normative account of belief then gives us a norm which is too weak to explain transparency, because transparency ensures that I cannot answer *whether p is true* in the positive and fail to believe that *p*. And here a similar diagnostic point as that made regarding the teleological account is available: the error is in relating transparency to something we do as believers—it is our commitment to the putative norm of belief which is said to generate transparency. But if that is right, there ought to be more (or any) scope for our doxastic deliberation to not be governed exclusively by considerations of truth. Normative accounts then also do not appreciate the passivity on the agential side: we *find ourselves* unable to deliberately form beliefs *but by* responding to considerations of truth.

As I noted earlier, there are moves the normativist might make in light of this objection. For example, she might argue that doxastic deliberation occurs within the context of inquiry which bolsters the permissibility norm, or that the permissibility norm directs one to the only permissible doxastic attitude one can have towards *p*, and so on. I have argued elsewhere that none of these amendments to the normativist's position help in an explanation of transparency, and so I direct the reader to those arguments (Noordhof and Sullivan-Bissett m.s.). For now, I hope I have said enough about the shortcomings of the two accounts which represent the current orthodoxy for explaining transparency, that the reader will be curious as to whether we can do better. I think that we can.

## 4 A motivational account of belief

We have seen that, at the very least, accounts of belief which understand it as necessarily connected to truth result in taking transparency to be an agential phenomenon, and in so doing, face problems. I suggest then that we sever the connection between belief and truth, and instead adopt a motivational account of belief. That is, what it takes for something to be a belief should be cashed out in terms of its motivational role. Recently philosophers have gotten a fix on the motivational role of belief by focusing on what happens when beliefs are true, given certain desires. Belief then has been understood as a state which interacts with desires to cause action such that when beliefs have true contents, *successful* action results. Neil Van Leeuwen refers to this as the ‘standard way’ of characterizing the motivational role of belief (Van Leeuwen 2009: p. 229). This is indeed a very standard view, and found in several other places.<sup>14</sup> Less standard is the claim that this motivational role is distinctive of belief, and it is a claim of this kind which I think constitutes the best account of the nature of belief. I outline and defend it briefly here.

Velleman takes the motivational approach as claiming that ‘all that’s necessary for an attitude to qualify as a belief is that it disposes the subject to behave in ways that would promote the satisfaction of his desires if its content were true’ (Velleman 2000: p. 255). There are good reasons to think that this characterization is extensionally inadequate, a subject may have ‘perceptions that *p*, desires that *p*, or questions as to whether *p*’ which may dispose him to form beliefs that *p*, and they may dispose him to behave such that his desires would be satisfied if *p* were true (O’Brien 2005: p. 55). These states are not beliefs, and so Velleman’s condition does not carve finely enough. Given this, Lucy O’Brien suggests the following amendment: for a state to be a belief, all that is necessary is that it ‘*by itself*, and relative to a fixed background of desires, disposes the subject to behave in ways that would promote the satisfaction of his desires if its content were true’ (O’Brien 2005: p. 56). I take up O’Brien’s characterization of belief’s motivational role.

What makes something a *belief* then, is not its being truth directed (or any nearby claim), but whether or not it has some motivational role.<sup>15</sup> If you do not find the condition I have put on belief appealing, feel free to replace it with your favourite motivational role of belief. Though I am not committed to any particular motivational role, I take O’Brien’s formulation to be a satisfactory one, but disagreement with me

<sup>14</sup> For example, David Armstrong understands Frank Ramsey’s claim that beliefs are maps by which we steer (Ramsey 1931) as suggesting that given ‘suitable dominant desires, then the belief-state will cooperate with the desire so that they are jointly responsible for the subject’s acting in a certain way’ (Armstrong 1973: p. 71). This view is also found in J. T. Whyte’s work on success semantics in which he claims that when our beliefs are true ‘we get what we want by acting on them’ (Whyte 1990: p. 149).

<sup>15</sup> What about beliefs which look motivationally inert? Consider my belief that *Neptune is very cold*. This might well be an attitude which will never motivate my action, am I thus forced to deny that this attitude is a *belief*? I am not, and that is because the motivational role I ascribe to belief is a conditional one—the attitude needs to be such that it *would* promote the satisfaction of my desires in certain contexts (e.g. in the context of a pub quiz), even if it does not, *in fact*, promote the satisfaction of my desires (because I avoid pub quizzes). I am grateful to an anonymous reviewer for raising this worry, and for their example.

on this point is unlikely to be damaging to my explanation of transparency which comes later. If the reader objects to distinguishing belief from other attitudes upon its motivational role I note that building belief's relationship to truth into the very nature of belief (as advocated by teleological and normative accounts) has, at best, given us problematic explanations of transparency, so it is worth seeing how far we can get with an alternative account of belief's nature.

One line of objection I should address, put most forcefully by Velleman (2000: pp. 255–256), and echoed by Shah and Velleman (2005: pp. 497–98), is that appeal to motivational role will not demarcate belief from other cognitive attitudes, which is precisely the point of a motivational account. Velleman claims that with respect to its motivational role, beliefs are no different from imaginings: a child imagining to be an elephant will be disposed to act in ways such that, if the content of the imagining were true, it, by itself, would promote the satisfaction of her desires (Velleman 2000: pp. 255–256). Similar claims might be made about some cases of acceptance.<sup>16</sup> However, this objection does not work, because the role that imagining (and accepting) can play in motivating action is context-dependent such that there will be cases in which the roles played by these attitudes and the role played by belief come apart. Once we enumerate these contexts, we can distinguish belief from other attitudes on the basis of its motivational role, and we can do this for any putative case of an attitude having the same motivational role as belief. Technically then, we should understand the motivational role of belief as being indexed *across contexts*.<sup>17</sup>

Can this account of belief explain transparency? We might say that, given belief's motivational role, true beliefs are going to be more useful than false beliefs—perhaps motivational theorists can appeal to this fact in an explanation of transparency. As Jerome Dokic and Pascal Engel in their discussion of Ramsey put the point:

a false belief can only be useful locally and by accident. A false belief cannot properly guide our behaviour in every possible situation. In contrast, truth systematically promotes the success of action, on a large scale. Truth offers a *guarantee* of success that falsity cannot hope to deliver. (Dokic and Engel 2002: p. 48)

Consider then the following claims: beliefs are those attitudes which have some particular motivational role (as above), and true beliefs guide our actions more successfully than false beliefs (where success is cashed out in terms of desire satisfaction). Given

<sup>16</sup> We might also consider a flip-side to Velleman's objection. Just as imagining that  $p$  seems to play the motivational role of believing that  $p$  in this context, it looks like the belief that *not-p* ceases to play its motivational role. I note two things here: first it might be that my belief that *I am not an elephant* might lose none of its motivational power after all, but that my belief that *I am pretending to be an elephant* takes precedent in the context of pretence. However, one might worry that importing a belief here makes childhood pretence look depressingly adult-like (Velleman 2000: pp. 257–258). In which case the second thing to say is that none of the actions I engage in whilst imagining that *I am an elephant* (e.g. waving my arm around) are inconsistent with my belief that *I am not an elephant*. And so it is not the case that my belief that *I am not an elephant* ceases to play its motivational role. Thanks to an anonymous reviewer for this take on Velleman's objection.

<sup>17</sup> Such a response is endorsed by Kathrin Glüer and Åsa Wikforss (2013: pp. 143–145), Paul Noordhof (2001: p. 253), and O'Brien (2005: p. 59).

these claims, our deliberative belief formation being characterized by transparency, being transparent to considerations of truth, is good news for believers. If my deliberation over *whether to believe that p* was not transparent to *whether p is true*, I might end up with more false beliefs, and thus, I may find myself engaging in less successful, and fewer successful, actions than I currently engage in. Perhaps then, we have found an explanation of transparency—when I consider the question *whether to believe that p*, my deliberation is transparent to the question *whether p is true* because I know that having true beliefs would be more likely to bring about successful action.

However, such an explanation runs into the same kind of problems as the ones I raised for teleological and normative accounts. And it does so in virtue of its framing transparency as an agential phenomenon. The explanation which goes via the nature of belief as a motivational state just sketched implies that I have control over the considerations that my deliberative belief formation is sensitive to. But if this were true then I would be able to answer the question *whether p is true* in the positive, without *thereby* believing that *p* (if I did not care for desire satisfaction, for instance). Once again then, the account under consideration gives us a gap between *whether p is true* and *whether to believe that p*. But there is no such gap in doxastic deliberation, truth is not an optional end. A motivational account of belief does not have the resources to explain why this is the case. I will not defend the motivational account of belief any further, since, however the details are filled in, it is not here that we will find an explanation of transparency.

Despite this explanatory shortfall, belief being an attitude with a given motivational profile *is* nevertheless important to an explanation of transparency, or so I will argue. The motivational account gets something right in its divorcing the nature of belief from truth, and it is by supplementing this account with one concerning the nature of *our* beliefs, that we can find an explanation of transparency. It is the particular circumstances in which *our* beliefs are produced which can explain this feature of their formation.

## 5 Supplementing the motivational account

Here I give my preferred explanation of transparency which comes from supplementing the motivational account of belief sketched above, with an account of actual world beliefs. The strategy then is to take a lesson from concerns about teleological and normative accounts—that casting belief as constitutively connected to truth is problematic for explanations of transparency—and thus sever the connection between belief and truth in an account of the nature of belief. Then, I add to this an account of *actual world* beliefs, i.e., those beliefs of humans. These actual beliefs, in virtue of being beliefs, will meet the motivational condition, but they will also have some additional features in virtue of the particular circumstances in which they are formed. It is these additional features which are key to explaining transparency. We need not, as others have, seek to explain transparency as something stemming from the nature of deliberation over what to believe. The resulting account then will move away both from framing transparency as a product of something believers have a

role in producing, and as something which is necessary to deliberation over what to believe.

What follows then is an account of *actual beliefs*, which I will label as such so as not to confuse it with the motivational account of belief's nature sketched above. I will also refer to the conjunction of my motivational account of belief and my biological account of actual beliefs as my *position*. In order to describe my account of actual beliefs, I need to lay some definitional ground. Firstly, I adopt an etiological account of biological function, which I have defended elsewhere (Sullivan-Bissett 2016). The details of the account do not matter here, so to simplify, for a trait to have a proper function it needs to possess one of a set of properties (cashed out in terms of an appropriate range of gene expression), and stand in a causal historical relationship to a token trait which possessed certain properties and was selectively successful.<sup>18</sup>

My account of function is influenced by Ruth Millikan's and so I will use some of her terms here which do not distinguish between our respective accounts (see Sullivan-Bissett 2016 for discussion of our differences). When we ascribe to a trait a *proper function*, we ascribe to it a function that was performed by its ancestors, which 'helped account for the proliferation of the genes responsible for it' and thus also 'helped account for its own existence' (Millikan 1989: p. 289). The heart has the proper function of pumping blood, because it originated as a reproduction of an ancestral heart, which performed the function of pumping blood in the past, and it is this which accounts for the proliferation of the genes responsible for the heart, and therefore the existence of it.

To say of a device that it has a *relational proper function* is to say of it that 'its function [is to] produce something that bears a specific relation to something else' (Millikan 1984: p. 39). Relational proper functions are the effects that have helped account for the selection of the producing mechanism (Millikan 1984: p. 26). To take a toy example; consider the chameleon's mechanisms for changing the skin pigmentation; they have the relational proper function of producing a skin pattern which matches the chameleon's environment (Millikan 1984: p. 39). The devices produced by such mechanisms are *adapted devices*, and that which the adapted device is adapted to is the *adaptor* (Millikan 1984: p. 40). The produced skin pattern of the chameleon in this instance is the adapted device, with the adaptor being the leaf which the device is produced in relation to.

Finally, I will use the term *Normal* (note the capitalization) to designate a normative historical (not statistical) sense of normality. For example, sperm Normally fertilize ova, but they do not normally fertilize ova ('Most never find an ovum and have to call it quits' (Millikan 1984: p. 34)).

<sup>18</sup> More fully: A token trait  $t$  is a member of the functional kind  $K$  and has  $F$ -ing as its selected function if and only if it: (a) Possesses one of a set  $S$  of intrinsic structural properties,  $\{p^1, p^2, p^3, \dots\}$ , and (b) Stands in a causal historical relationship to at least one token trait which (i) possessed one of a set  $S$  of intrinsic structural properties,  $\{p^1, p^2, p^3, \dots\}$ , (ii) possessed one of a set  $S^*$  of properties responsible for that token trait  $F$ -ing,  $\{q^1, q^2, q^3, \dots\}$ , and (iii) was selected for  $F$ -ing (Sullivan-Bissett 2016). As I noted, the details do not matter here, any historical account of biological function can be plugged into my account of actual beliefs and explanation of transparency.

## 5.1 Actual beliefs, truth, and function

With the relevant terms defined, I now offer my functional claim about actual beliefs. Recall that I am building a biological account of actual beliefs onto a motivational account of the nature of belief *simpliciter* (i.e. belief across worlds).<sup>19</sup> A belief is a state which plays a certain motivational role, and a biological account of this kind of state as it occurs in the actual world can do the explanatory work called for by transparency. Such an account will ascribe a function to our mechanisms for belief-production, which will provide an explanation of the constraints under which they operate.<sup>20</sup>

The mechanisms which produce beliefs in us have the relational proper function of producing true beliefs.<sup>21</sup> When mechanisms produce devices in accordance with their relational proper function, they produce devices which are *adapted to* something else. Our mechanisms for belief-production Normally produce devices which are *adapted to* the way things are in our environment, that is, they Normally produce beliefs with true contents. Given the biological advantages enjoyed by a creature with mechanisms which produce true beliefs, the mechanisms' producing such items, causally explains why we have these mechanisms now. Such mechanisms have proliferated because—given the essential nature of belief (its motivational role)—creatures with true beliefs will be more likely to survive because their actions will be successful, given a certain ecology (like the one humans evolved in). The Normal explanation for the performance of this function is that the creature acts on the true belief in such a way as to satisfy its desires.<sup>22</sup>

<sup>19</sup> Perhaps a little reflection on the relationship between these two steps is needed. We might wonder whether biological proper function accounts of actual beliefs *presuppose* motivational accounts of belief, or, weaker, whether we should *expect* motivational accounts of belief to be supplemented by biological proper function accounts of actual beliefs. As to the first point, I think biological proper function accounts of actual beliefs need not presuppose motivational accounts of belief. As to the second point, my view is that we should expect motivational accounts of belief to be supplemented by biological proper function accounts of actual beliefs *insofar as one wants to explain transparency*, but not insofar as one wants to demarcate belief. Motivational accounts of belief cannot do the former work on their own, but they can, on their own, do the latter work. I am grateful to an anonymous reviewer for requesting clarity here.

<sup>20</sup> What if there were attitudes produced by these mechanisms which did not have the motivational role identified as distinctive of belief? In that case the machinery I have suggested is in play would not be quite right. Rather, we would have some biological mechanism which produces states—sometimes *beliefs*, sometimes not. But then everything would proceed as before. The scope of my position would be restricted to those mechanisms in circumstances in which they produce *beliefs*, and the position would explain why those items exhibit transparency. Thanks to an anonymous reviewer for raising this point.

<sup>21</sup> Millikan has made some of the claims about belief which I do, she has not however, put them to work in the same way I do. She claims, for example, that it is a 'function of our belief-fixing systems to fix true beliefs' (Millikan 1995a: p. 177, 1995b: pp. 243–244). She ascribes to belief several functions, including '[participating] in inferences in such a manner as to help produce fulfillment of desires' and '[participating] in inferences to yield other [true] beliefs' (Millikan 1995c: p. 71). In this paper I ascribe only one function to belief. Not because I think that Millikan is wrong in her multiple function ascriptions, but because I am interested in how much work an appeal to biological function can do in explaining transparency. My relatively limited function ascriptions then should be seen to reflect my relatively limited project.

<sup>22</sup> Elsewhere I ascribe a second function to our mechanisms of belief production—the function of producing *organizing* beliefs—which are useful to the believer, but not as an approximation to truth. Such a function is performed in circumstances abnormal for the production of true beliefs, in circumstances in which

I note a concern that might be had at this point: if the transition between the deliberative questions is explained by the contingent fact that believing truly is (usually) adaptive, then would it not be that the deliberative question would collapse into the *practical question* ‘would believing  $p$  be adaptive/useful?’ However, the problem with this line of argument is that it suggests that the transition between the questions is agential – that it can be put down to something we *do*. This is precisely the kind of picture I reject. There is a difference between there being a reason for the transition between the questions (belief’s being a motivational state and our biological history setting us up in such-and-such a way), and the agent having that story *as a reason*. My claim here is that because believing truly, in circumstances Normal for the production of true beliefs, is adaptive, we are set up such that our deliberative belief formation is transparent to truth. But I am not saying that we, as believers, recognize this *as a reason* to focus on truth. Our relationship to the structure of deliberative belief formation is much more passive than this—indeed, this is one of the take home messages of the paper.

At this point the reader could be forgiven for thinking that my account of actual beliefs is teleological and thus vulnerable to worries I raised earlier, perhaps it looks like our mechanisms for belief-production are *aiming* at truth. However, on my account, the only sense in which actual beliefs have an *aim* is the sense in which they have a biological proper function. My account of actual beliefs captures the agent’s relatively passive response to truth suggestive of biological function, rather than settled by anything at which the agent can be said to be aimed. My account of actual beliefs is teleological but rooted in biology, and does not include an appeal to actual beliefs being *aimed* at truth in anything but a sub-intentional way. This means that I do not face the dilemma raised earlier, since I explain transparency by appeal to one of the biological functions of the mechanisms which produce actual beliefs, not by appeal to any aim of the agent. There is thus no problem for my account of actual beliefs posed by the exclusivity of truth considerations in deliberation over what to believe, since I do not explain transparency by appeal to an aim, the moderation of one’s activities by which ought to allow for weighing. The aims of action are set by desires and desires can be weighed against each other. So aims in the context of action theory are weighable, or so it seems reasonable to say. Of course there is a sense of *aim*—that provided by my account of actual beliefs—where aim-talk is just another way of capturing function-talk. But to adopt this line is just to concede that talk of aims is not functioning at the level of action as the teleologist has in mind.

---

Footnote 22 continued

truth-directed mechanisms do not facilitate the effective functioning of the believer. The work here roughly tracks the distinction drawn by Ryan McKay and Daniel Dennett (2009) between two types of misbelief: those resulting from malfunction, and those resulting from Normal functioning. Misbeliefs of the second type are the result of mechanisms of belief-production functioning as they should. It is misbeliefs of this type which, I argue, are produced by our mechanisms of belief production performing the function of producing organizing beliefs. I recognize then that natural selection has built our mechanisms for belief-production with more than tracking truth in mind, that this is so is demonstrated by a quick tour of beliefs arising from self-deception and other doxastic biases (see relevant sections of Haselton and Nettle 2006 for an overview of doxastic biases which may be favoured by natural selection). However, since this second function ascription is not relevant for an explanation of transparency, I do not discuss it here (see Sullivan-Bissett (2017b)).

Nor is my motivational account of belief vulnerable to Shah's teleologist's dilemma, since my descriptive condition on something's being a belief does not go via its connection to truth. On my view, an attitude is a belief if it meets some motivational condition, and not if it has some truth-regulatory feature. I do not then face the unhappy task of stipulating the level of truth-regulation required for something to be a belief, such that it captures all those states we want to capture under that label, and it also explains exclusivity to truth in some cases of belief formation.

## 5.2 The biological function of actual beliefs and transparency

I now show how my account of actual beliefs against a background of a motivational account of belief simpliciter, can explain transparency. The explanation I offer appeals to those causal facts about actual belief formation which obtain in virtue of the natural selection of mechanisms which produce beliefs with true contents. As a proponent of an etiological account of biological function, I claim that the fact that a device has a certain function is an explanation for that device's presence today.

When a subject poses to herself the question *whether to believe that p*, she attends to the question *whether p is true*. To demonstrate the structure of my explanation, we can usefully ask two questions about the presence of transparency in doxastic deliberation. First, we can ask a *how* question: how is there transparency in doxastic deliberation? This question demands an answer providing details of the mechanisms responsible for causing transparency. The explanatory burden here is to give an account of why, when a subject deliberates over *whether to believe that p*, she moves immediately and inescapably to considering the question *whether p is true*. The grounds of transparency—of our moving thus—can be given by appeal to certain neurological structures.

In functional terms: transparency is achieved by certain causal facts which hold for our mechanisms for belief-production. Such causal facts make it such that whenever our mechanisms for belief-production Normally perform their relational proper function and produce a true belief, the adaptor in each case is the way the environment is. This is understood at the agent-level as epistemic considerations pertaining to the truth of *p*. Instances of epistemic considerations settling whether to believe are instances of believers immediately and inescapably attending to the adaptor (the environment) for the adapted device (the belief).

This explanation of how transparency is realized is the same kind of explanation given by the teleological and normative accounts discussed above. All three explanations appeal to something which grounds our moving from the question *whether to believe that p* to the question *whether p is true*. On the teleological account, this is explained by appeal to the aim one adopts in posing the deliberative question, on the normative account it is explained by appeal to one's being guided by the norm governing belief. I argued that these explanations are problematic insofar as the aim and norm are not at the required strength to explain transparency (Sects. 2.1, 3.1). An appeal to biological function does not face this problem, since the move from one question to another is not achieved by anything the agent *does*, but is rather secured by the biological history of our mechanisms for belief-production. The agent's role is a

passive one, and thus there is no question of answering *whether p is true* in the positive, and withholding belief because of one's other aims, or because one recognizes that belief is merely permissible, and not obligatory.

Second, we can ask a *why* question: why do we move from the question *whether to believe that p* to the question *whether p is true*? What explains why the causal facts obtain, which ground our moving between the questions in this way? This question demands an answer which explains why transparency characterizes our deliberation and not something else, realized by a different set of neurological structures. This explanation will tell us why *these particular causal facts* obtain.

Here we can appeal to natural selection, which selected for the particular neurological structures which ground the move from the question *whether to believe that p* to the question *whether p is true*. Given that our mechanisms for belief-production have the relational proper function of producing devices with true contents, the story for their selection includes their Normally producing true beliefs. In the deliberative case our cognitive architecture is arranged such that when we deliberate as to *whether to believe that p*, we, at the agent-level, are only sensitive to the adaptor (our environment) because this makes the adapted device (the resulting belief) more likely to perform its derived proper function of being true. The neurological structures which secure transparency have been selected for their role in producing true beliefs.

The teleological and normative accounts answer the *why* question by appeal to the very nature of belief. Why is it that we move from *whether to believe that p* to *whether p is true*? Because it is part of the very nature of belief that it is aimed at truth, or governed by a norm of truth. In contrast to these accounts then, my claim is that we have a pre-commitment to truth when we deliberate, this pre-commitment is secured by causal facts which constitute the cause of transparency, and my account of actual beliefs can also explain why these particular causal facts obtain.

My account of actual beliefs leaves it open that there might have been other explanations for why these particular causal facts obtain, and hence that transparency is not essentially to do with biological function. For example, God might make it the case that actual beliefs are transparent to truth considerations. He might ensure that we move from the question *whether to believe that p* to the question *whether p is true*, by instantiating certain neurological structures. Leaving this possibility open is all to the good: I embrace the consequence of my account that transparency is not a necessary feature of deliberative belief formation, nor is the way in which it is contingently realized the only way that it might be realized. Other world deliberative beliefs might too be characterized by transparency, without this being explained by biological function. Transparency as exhibited by actual world doxastic deliberation, though, is explainable by appeal to the biological function of the mechanisms which produce beliefs in us.

In the next two sub-sections I will outline and respond to an objection to my explanation of transparency, as well as the supposed virtues of a nearby alternative position which I prefer not to endorse.

### 5.3 Contingency is unacceptable

My explanation appeals to natural selection to explain why the causal structures which realize transparency are present. This means that transparency comes out as a contingent feature of the deliberative beliefs of some believers (i.e., those believers with a biological history), not as a necessary feature of deliberation over what to believe. An objection might be that this consequence of my overall position is unacceptable, and is a reason to reject it. Perhaps it is to be taken as part of the explanandum of transparency that it holds as a matter of metaphysical necessity. This might be taken as something which ideally would be explained, but equally, something which ought not to be abandoned because of failures to explain it.

Here then, I should justify my moderation of the modal strength of transparency. As I argued earlier, other accounts of the nature of belief offer problematic explanations of transparency (at any strength). Such accounts, I suggested, make the mistake of locating the role of truth in belief formation at the intentional level, as something which a subject can aim at, or express her commitment to via being guided by some norm.

Our explanandum is a descriptive fact about belief formation, and I see no reason to think that that descriptive fact is one which is co-instantiated with every instance of deliberative belief across worlds. What needs explaining is why when *we* deliberate over *whether to believe that p*, that question collapses into *whether p is true*, which is why I move from talking about the nature of *belief*, to talking about the nature of *actual beliefs*. Why should we think that we ought to build transparency, a feature of *actual* doxastic deliberation into a claim about the nature of such deliberation *simpliciter*? Our reason cannot be that belief is constitutively aim-governed, that will not buy transparency, as I have argued (Sect. 2.1). Nor can it be that belief is constitutively norm-governed, since the only plausible formulation of the norm is one which is not strong enough to generate transparency, as I have argued (Sect. 3.1). Teleological and normative accounts then do not necessitate transparency for beliefs formed in the actual world by human believers, let alone for beliefs formed in other possible worlds.

It might be that transparency being taken to have this strong modal status is a function of explanations of it, rather than anything about the phenomenon itself. At least in the case of the teleological account, the explanation of transparency goes via an appeal to the kind of thing belief necessarily is, which is then said to entail the presence of transparency. This has the result that just as the nature of belief goes across worlds, so too does transparency. But given that it is, at the very least, not clear that the teleological account can explain transparency, why keep hold of one of the proposed explanation's entailments?

Perhaps this is too quick. We might be able to build a case for the strong modal status of transparency independently of the theoretical status of accounts which entail that claim. A natural way of arguing for a necessity thesis about transparency is by appeal to the concept of belief: we might say that it is part of the concept of belief that it is an attitude which exhibits transparency, since it is an attitude which is governed 'both normatively and descriptively, by the standard of truth' (Shah and Velleman 2005: p. 499). And so, deliberation over whether to form an attitude that *p* which does not exhibit transparency is not deliberation over whether to form a *belief*. One thing

to say here is that building the connection to truth into the nature of belief from the start will be friendly to this kind of conceptual claim, but we have already seen that these accounts run into problems when it comes to doing the explanatory work. So I think it is wise to be open to an alternative position, that of understanding the nature of belief as divorced from truth, and then looking to the particular circumstances of belief formation in the actual world. An apparent cost of this alternative is that the necessity thesis is not honoured, and so now we are presented with a choice: pick the accounts of belief which give us problematic explanations of transparency, but honour it as modally strong, where the grounds for understanding it at this strength are given by appeal to a conceptual claim about belief. Or, pick an account of actual beliefs coupled with a motivational account of the nature of belief. This position can explain transparency, but does not honour the strong modal status, and so loses points with the advocates for this conceptual claim. I think that the debate bottoms out at this point.

I also have a principled reason, from my motivational account of belief, to disagree with the conceptual claim Shah and Velleman make which props up transparency as modally strong. Shah and Velleman claim that transparency is distinctive of belief, and it is by appeal to it, that we can distinguish belief from supposition (Shah and Velleman 2005: p. 499). But the motivation for this claim is lost given my motivational account of belief—there is no reason to add to this with respect to providing a demarcating criterion of belief. The motivational role specified earlier captures an important generality across believers, with the link to truth given by features such as transparency as an add-on, which are thus contingent features of belief. To take transparency to be a necessary feature of deliberation over what to believe would just be to project the particular circumstances of our biological heritage into a modal claim.

Given that there are no good reasons to be committed to the necessity claim about Transparency, a position like mine which casts it as contingent should not be objected to on the basis of not honouring the necessity claim. If my position combining belief as a necessarily motivational state with a biological story of actual beliefs is able to do the explanatory work with respect to transparency, then the fact that it does not honour a thesis for which there are no good grounds is surely not a strike against it.

#### 5.4 A stronger function account

I might be advised that the problem outlined above can be solved by being more serious about function, that is, by claiming that it is essential to the nature of belief that it is produced by mechanisms with the biological function of producing true beliefs. Such an account would cast my functional claim about actual beliefs as a claim about the kind of thing belief necessarily is (and not merely a claim about the etiology of actual world beliefs). This means that I would not need the motivational role specified earlier as necessary and sufficient for something's being a belief. Rather, so this view would go, if an attitude is not produced by mechanisms with the biological function of producing true beliefs, whatever that attitude is, it is not a belief. The suggestion then is to take the function claim about actual beliefs to a different level of explanation—to embed that functional claim into one's account of belief simpliciter. If I were to adopt this account of the nature of belief, there would be no problem arising from casting transparency

as contingent. If transparency is to be explained by the biological function of beliefs, and if all beliefs have that biological function as a matter of necessity, transparency will also hold as a matter of necessity.<sup>23</sup>

Though this stronger function account would not be vulnerable to the objection I raised above, we have seen that my position has the resources to diffuse it. So, this stronger function account is not required, and additionally, it raises a particularly difficult problem, which my preferred position does not face. In discussing a component of Velleman's teleological view, Glüer and Wikforss raise an objection to an account of belief which has as a necessary and sufficient feature of belief the function of the mechanisms which produce it. This view would deny beliefs to an individual without the appropriate history, that is, a history which would bestow upon its mechanisms of belief-production a proper function (Glüer and Wikforss 2013: p. 145). Donald Davidson's Swampman makes the point: Imagine that I, and a dead tree in a swamp, get struck by lightning. My body is 'reduced to its elements', whilst the tree is turned into a molecule for molecule identical version of me (Davidson 1987: p. 443). This replica continues to write philosophy papers, expresses desires to go for 5 km runs, and reports on deliberating about her partner's (in)fidelity. On the strong function account, my replica is ruled out of having beliefs because of her (lack of) history. Perhaps a proponent of the strong functional account would not balk at this consequence of her view, if she is a historical function theorist then she is accustomed to putative counterexamples which work by removing history and insisting on the presence of some feature nevertheless. So perhaps a strong function theorist would be happy with the consequence that a creature like my Swamp replica could not have beliefs, governed by transparency or otherwise.

However, this is at the very least, it seems to me, a theoretical cost: presumably the strong function theorist would grant the metaphysical possibility of a creature for whom the same causal facts held as do for us. If such a creature were physically and functionally identical to a biological creature, denying that such a creature had beliefs is surely best avoided. My position does not face this problem since I take the functional claim to be a contingent one, insofar as it is a claim about actual world

<sup>23</sup> Something close to this view can be found in a recent paper by Kate Nolfi (2015). She claims that 'what makes belief the kind of mental attitude that it is, is just that beliefs are supposed to fulfill a distinctive proper function' (2015: 197). The function she identifies is parsed in motivational terms: 'the proper function of belief is to inform our decisions to act by serving as a kind of map of the way things are so that we achieve whichever ends our actions aim to achieve' (Nolfi 2015: p. 197). She claims that the notion of 'proper function' she is working with is *biological* proper function (Nolfi 2015: p. 203, n. 27). Her motivational role of belief then is indexed to the biological function of belief, which itself is taken to be *necessary to belief*. Nolfi claims that her account is intended to be compatible with any account of biological function. However, different notions of biological function plugged into Nolfi's account will have consequences for that account. For example, if she adopts a historical account of function, then her view will also have the consequence that creatures without the right history cannot have beliefs. If she adopts a forward-looking account of function, then she will not be able to ground the notion of a malfunctioning belief, and she will not be able to explain why belief has the function that it does (only historical accounts have been thought able to explain the existence of the functionally characterized item, and accommodate such items malfunctioning) (Millikan 1989; Sullivan-Bissett 2016).

beliefs, not a claim about belief simpliciter.<sup>24</sup> Providing the state in question has the appropriate motivational role, that state is a belief. My spontaneously coalescing replica can *believe* that transparency is best explained by appeal to biological function, just as I do.

## 6 Conclusions

I have argued that two kinds of explanation for transparency currently on offer in the literature are problematic. The truth-aim teleological account of belief faces a dilemma: either there is no aim of belief, in which case it is not legitimate to appeal to one in explaining transparency, or there is an aim, but it is not at the required strength to explain transparency. In addition, Shah's teleologist's dilemma reveals a tension between a descriptive claim about truth-regulation being distinctive of belief, and that claim explaining transparency—it cannot do both. The best version of the normative account of belief is one which casts the norm governing belief in terms of permissibility. However, once the normativist moves to this weaker norm, it may well be too weak to explain transparency. I suggested that these accounts run into trouble because they offer explanations of transparency which involve something the believer does. I then outlined my preferred account of the nature of belief which went via its motivational role. After showing that this account on its own cannot explain transparency, I supplemented it with an account of actual world beliefs. This account had it that our mechanisms for belief-production have the relational proper function of producing true beliefs. This is an account of actual beliefs which reflects the relationship of passivity we have with transparency. I showed that this account of actual beliefs can explain transparency, responded to an objection to my explanation, and justified my not opting for a nearby alternative position. In conclusion then, transparency can be explained by adopting a motivational account of the nature of belief, and appealing to a biological function of our mechanisms for actual world belief-production, and is thus a contingent feature of the deliberative belief formation of some believers.

**Acknowledgements** I acknowledge the support of a European Research Council Consolidator Grant (Grant agreement 616358). I am grateful to audiences at the Mind and Reason research group at the University of York, the Transparency in Self-Knowledge and Belief workshop at the University of Oviedo, the Early Career Mind Network Meeting at the University of Cambridge, and the Southern Normativity Group's Annual Conference at the University of Sussex. Thank you also to Craig French, Miriam McCormick, Paul Noordhof, Kathy Puddifoot, Kirk Surgener, Henry Taylor, Jane Tomlinson, and Jonathan Way for very helpful comments on earlier versions of this paper. Finally, thank you to two anonymous reviewers for this journal.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

<sup>24</sup> In this respect, my account of actual beliefs is similar to David Papineau's version of teleosemantics. Just as the 'essential core of teleosemantics' is a claim about states in the actual world (Papineau 2001: p. 286), so too is the biological function component of my position a claim about beliefs in the actual world.

## References

- Armstrong, D. M. (1973). *Belief, truth and knowledge*. London: Cambridge University Press.
- Bykvist, K., & Hattiangadi, A. (2007). Does thought imply ought? *Analysis*, 67(4), 277–285.
- Bykvist, K., & Hattiangadi, A. (2013). Belief, truth, and blindspots. In T. Chan (Ed.), *The aim of belief* (pp. 100–122). Oxford: Oxford University Press.
- Donald, D. (1987). Knowing one's own mind. *Proceedings and Addresses of the American Philosophical Association.*, 60(3), 441–458.
- Dokic, J., & Engel, P. (2002). *Frank Ramsey: Truth and success*. London: Routledge.
- Engel, P. (2007). Belief and normativity. *Disputatio*, 2(23), 179–204.
- Glüer, K., & Wikforss, Åsa. (2013). Aiming at truth: On the role of belief. *Teorema*, 42(3), 137–162.
- Haselton, M. G., & Nettle, D. (2006). The paranoid optimist: An Integrative evolutionary model of cognitive biases. *Personality and Social Psychology Review*, 10(1), 47–66.
- McCormick, M. S. (2015). *Believing against the evidence: Agency and the ethics of belief*. Abingdon: Routledge.
- McHugh, C. (2011). What do we aim at when we believe? *Dialectica*, 65(3), 369–392.
- McHugh, C. (2012a). Beliefs and aims. *Philosophical Studies*, 160(245), 425–439.
- McHugh, C. (2012b). The truth norm of belief. *Pacific Philosophical Quarterly*, 92, 8–30.
- McHugh, C. (2013a). Normativism and doxastic deliberation. *Analytic Philosophy*, 54(4), 447–465.
- McHugh, C. (2013b). The illusion of exclusivity. *European Journal of Philosophy*, 23(4), 1117–1136.
- McKay, R., & Dennett, D. (2009). The evolution of misbelief. *Behavioral and Brain Sciences*, 32, 493–561.
- Millikan, R. G. (1984). *Language, thought and other biological categories*. Cambridge: MIT Press.
- Millikan, R. G. (1989). In defense of proper functions. *Philosophy of Science*, 56(2), 288–302.
- Millikan, R. G. (Ed.). (1995a). Explanation in biopsychology. In *White queen psychology and other essays for alice* (pp. 171–192). Cambridge: MIT.
- Millikan, R. G. (Ed.). (1995b). Naturalist reflections on knowledge. In *White queen psychology and other essays for alice* (pp. 241–264). Cambridge: MIT.
- Millikan, R. G. (Ed.). (1995c). Thoughts without laws. In *White queen psychology and other essays for alice* (pp. 51–83). Cambridge: MIT.
- Moran, R. (1988). Making up your mind: Self-interpretation and self-constitution. *Ratio (New Series)*, 1, 135–151.
- Moran, R. (2001). *Authority and estrangement*. Princeton: Princeton University Press.
- Nolfi, K. (2015). How to be a normativist about the nature of belief. *Pacific Philosophy Quarterly*, 96, 181–204.
- Noordhof, P. (2001). Believe what you want. *Proceedings of the Aristotelian Society*, 101, 247–265.
- Noordhof, P., & Sullivan-Bissett, E. m.s Transparent failure of norms to keep up standards of belief.
- O'Brien, L. (2005). Imagination and the motivational view of belief. *Analysis*, 65(1), 55–62.
- Owens, D. J. (2003). Does belief have an aim? *Philosophical Studies*, 115(3), 283–305.
- Papineau, D. (2001). The status of teleosemantics, or how to stop worrying about swampman. *Australasian Journal of Philosophy*, 79(2), 279–289.
- Ramsey, F. P. (1931). *The foundations of mathematics and other logical essays*. London: Routledge.
- Shah, N. (2003). How truth governs belief. *The Philosophical Review*, 112(4), 447–482.
- Shah, N., & Velleman, J. D. (2005). Doxastic deliberation. *The Philosophical Review*, 11(4), 497–534.
- Steglich-Petersen, A. (2006). No norm needed: On the aim of belief. *The Philosophical Quarterly*, 56(225), 499–516.
- Steglich-Petersen, A. (2008). Does doxastic transparency support evidentialism? *Dialectica*, 62(4), 541–547.
- Steglich-Petersen, A. (2009). Weighing the aim of belief. *Philosophical Studies*, 145(3), 395–405.
- Steglich-Petersen, A. (2011). How to be a teleologist about epistemic reasons. In A. Reisner & A. Steglich-Petersen (Eds.), *Reasons for belief* (pp. 13–33). Cambridge: Cambridge University Press.
- Steglich-Petersen, A. (2017). Weighing the aim of belief again. *Logos and Episteme*, 3(1), 141–146.
- Sullivan-Bissett, E. (2016). Malfunction Defended. *Synthese.*, doi:10.1007/s11229-016-1062-8.
- Sullivan-Bissett, E. (2017a). Aims and exclusivity. *European Journal of Philosophy*. doi:10.1111/ejop.12183.
- Sullivan-Bissett, E. (2017b). Biological function and epistemic normativity. *Philosophical Explorations*, 20(1), 94–110.

- 
- Sullivan-Bissett, E. (forthcoming). Transparency defended. In *Syndicate Philosophy. Symposium on Miriam McCormick's Believing Against the Evidence: Agency and the Ethics of Belief*.
- Sullivan-Bissett, E., & Bortolotti, L. (forthcoming). Fictional persuasion, transparency, and the aim of belief. In E. Sullivan-Bissett, H. Bradley, & P. Noordhof (Eds.), *Art and belief*. Oxford: Oxford University Press.
- Sullivan-Bissett, E., & Noordhof, P. (2013). A defence of Owens exclusivity objection to beliefs having Aims. *Philosophical Studies*, 163(2), 453–457.
- Sullivan-Bissett, E., & Noordhof, P. (2017). Another defence of Owens's exclusivity objection to beliefs having aims. *Logos and Episteme*, 3(1), 147–153.
- Van Leeuwen, Neil. (2009). The motivational role of belief. *Philosophical Papers*, 38(2), 219–246.
- Velleman, D. J. (2000). *The possibility of practical reason*. Oxford: Oxford University Press.
- Wedgwood, R. (2002). The aim of belief. *Philosophical Perspectives*, 16, 267–297.
- Whiting, D. (2010). Should I believe the truth? *Dialectica*, 64(2), 213–224.
- Whyte, J. T. (1990). Success semantics. *Analysis*, 50(3), 149–157.