

**teorema**

Vol. XXXV/1, 2016, pp. 13-34

ISSN: 0210-1602

[BIBLID 0210-1602 (2016) 35:1; pp. 13-34]

## Can There Be Belief Without Evidence?\*

Sam Wilkinson

### RESUMEN

Presento una concepción de la creencia (denominada “solo descendente” [*downstream only*]) de acuerdo con la cual puede haber creencia sin evidencia. Defiendo esta tesis en contra de dos influyentes objeciones. La primera objeción está fundamentada en la intuición de que la creencia “apunta hacia la verdad”. La segunda de ellas dice que, puesto que hay estados que intuitivamente no son creencias, pero que explican la acción de la misma manera, necesitamos consideraciones de tipo ascendente [*upstream*] para distinguir entre esas creencias y los estados que no son creencias.

PALABRAS CLAVE: *creencia, normas, evidencia, acción, alief.*

### ABSTRACT

I present a view of belief (called the “downstream-only view”) according to which there can be belief without evidence. I defend it against two influential objections. The first objection is grounded in the intuition that belief “aims at truth”. The second objection is that, since there are states that intuitively aren’t beliefs but that explain action in the same way, we need upstream considerations to tell these belief and non-belief states apart.

KEYWORDS: *Belief, Norms, Evidence, Action, Alief*

### I. INTRODUCTION

It is important from the outset to get clear on the nature of the title question. The question is not whether human beings are psychologically capable of belief without evidence, or whether there are actual instances of people believing without evidence. The question concerns whether the very concept of belief allows there to be such a thing as belief without evidence.

Many theorists take there to be an essential relation between belief and evidence. Thus, cases that may look to us like instances of belief, but which lack the appropriate relations to evidence, fail to actually be instances of belief. Contrary to this popular view, I will present a view of

belief (called the “downstream-only view”, for reasons that will become clear) according to which there can be belief without evidence. I will then defend it against two influential objections.

## II. REVISIONISM AND CONSERVATISM ABOUT BELIEF

As a preliminary, it is very important to distinguish between *revisionism* and *conservatism* about belief. A conservative will take as a desideratum for a characterization of belief that it be in line with our actual practices of belief attribution. A revisionist, on the other hand, will deny this. Rather, belief, *in the relevant theoretical sense*, is something with a clearly definable nature that is independent of our practices of belief attribution. As a result, there may be cases where we attribute beliefs, but the subject doesn’t really believe what our attributions suggest.

### II.1. *An Illustration of Revisionism*

A very strong, and explicit, example of revisionism about belief is to be found in the work of Ruth Barcan Marcus (most explicitly in her 1995 paper). She claims that belief, in the relevant sense of central theoretical importance, is a relation, not to sentences (or propositions conceived as sentence-like), but to possible states of affairs, and, as a result, she claims that one cannot believe the impossible.<sup>1</sup> This is deeply revisionist since we commonly can ascribe impossible beliefs to others. For example, mistaken beliefs about identity are often ascribed as, “He believes that Batman is not Bruce Wayne”, and it is impossible (true in no possible worlds) that something should fail to be self-identical.<sup>2</sup> But according to Marcus, no one can actually believe this, since you cannot stand in the belief relation to impossible states of affairs. Marcus instead explains away the mere appearance that people can believe the impossible.

We will not need to take a stance on this tangential debate about believing the impossible. However, it is a nice illustration of what I mean by revisionism, namely, of our belief attributions coming apart from actual instances of belief in some theoretically important sense.

### II.2. *An Illustration of Conservatism*

Lisa Bortolotti’s book, *Delusions and Other Irrational Beliefs* (2009), is devoted to defending the claim that delusions are beliefs. Her tactic is to present those who deny that delusions are beliefs (so-called “anti-doxasticists”) with a dilemma: if we are to deny belief-status to delusions,

then we are going to have to do the same for many states that, intuitively, we are happy to think of as beliefs. Among these “mental states that we are happy to regard as typical beliefs” [Bortolotti (2009), p. 57], she cites, in particular, the sort of biased hypothesis-testing we commonly get in scientific beliefs [Ibid., p. 148] and also the unrevisability of racist [Ibid. p.150] and religious [Ibid. p.152] beliefs. Note that this argument is based on the suggestion that denying belief-status to these non-pathological states, that we normally attribute as beliefs, is too great a cost. In other words, it takes as a desideratum for a theoretical characterization of belief that it should align itself with the beliefs that we attribute on a daily basis. A characterisation of beliefs that denies belief-status to delusions also has as a consequence that religious, scientific and racist beliefs are not “really” beliefs, and this is an unacceptable consequence. This is clearly conservative in the sense just mentioned, namely, since it takes our actual belief attributions as a guide for what is to actually count as belief.

This is an intuitively appealing strategy. However, note that the anti-doxasticist can counter such a strategy by affirming a revisionism about belief, and simply question that denying belief-status to, e.g., certain religious “beliefs”, is too great a cost. The anti-doxasticist might say, as Ruth Barcan Marcus has, that it may turn out that many of the times that we regularly attribute beliefs to others, we are wrong and they are not, strictly speaking, believing in the proper, revised, sense. Furthermore, one might think that this is a perfectly acceptable, or even attractive, consequence of a *philosophical* theory of belief. What is philosophy for, after all, if not to regiment and revise our pre-theoretical concepts? We may ascribe belief in a rough and ready way, but that doesn’t mean that belief is a rough and ready phenomenon.

### II.3. *Conservatism, Revisionism and Belief without Evidence*

Bortolotti’s work nicely shows that, given the continuity (in particular in terms of irrationality) between clinical delusions and non-pathological beliefs, if the anti-doxasticist position is going to be plausible, one has to be a revisionist about belief. The same applies to those who deny the possibility of belief without evidence. To view things from the other direction, so to speak, someone who wants to claim that there is belief without evidence, and does so on the grounds that we attribute various forms of evidence-less beliefs (religious beliefs, delusions etc.) is open to the same revisionist move. It is argumentatively more efficacious

to show that one can be a revisionist, and allow for belief without evidence. As it happens, the position I want to present is precisely that.

### III. BELIEF: A DOWNSTREAM-ONLY VIEW

In this section, I present the central view. It is that only “downstream considerations” are relevant in determining whether something is a belief or not. Given that relationships to evidence are upstream considerations, this means that a belief-like state can be without evidence and it could still be a belief, as long as it has the right consequences.

#### III.1. *Belief, Direction of Fit, and Maps*

What kinds of entities are at least candidate believers? One way of answering this question is by considering the teleological question: “What kinds of entities *need* something like belief?” One intuitive response is to say that organisms need information about the world, and hence are candidate believers. Not every organism is a believer, but the claim that organisms believe and, say, rocks don’t is at least a start. Then we might want to narrow this down a bit further to organisms that move. Entities that have no means of operating in the world have no need for an informational view of it. Then we might want to narrow it down further to organisms that move autonomously in the service of their goals. Animals navigate environmental features in the service of their goals, and presumably they do so by having some kind of information about the world, perhaps even something we could characterize as an informational perspective on the world.<sup>3</sup>

Frank Ramsey (1931) put belief precisely in these terms. He called beliefs “the maps whereby we steer”.<sup>4</sup> This metaphor of beliefs as being like maps that help us through the world in the service of our goals is extremely useful. In particular, it picks out an opposition between the phenomenon of belief (*viz.* the map) and the phenomenon of desire (*viz.* the goal). The former involves *taking* the world to be a certain way, whereas the latter involves *wanting* it to be a certain way. This picks out a fundamental distinction between information and motivation.

This opposition between belief (information) and desire (motivation) has sometimes been put in terms of “direction of fit” (which in turn is derived from speech act theory, where it was used to elucidate the difference between assertions and commands). The first use of the term “direction of fit” was by J.L. Austin (1962), and was not in fact used to

pick out this distinction at all.<sup>5</sup> His student, John Searle, picked up the terminology, and used it to pick out a distinction that he claims is best illustrated in Anscombe (who, in fact, did not use the “direction of fit” terminology). Searle’s merging of Austin’s terminology and Anscombe’s example has stuck, and has been picked up by various philosophers since. The example Searle (1979) takes from Anscombe is one that distinguishes a shopping list, from a detective’s record of what the man doing the shopping is buying:

...if the [shopping] list and the things that the man actually buys do not agree, and if this and this alone constitutes a mistake, then the mistake is not in the list but this man’s performance (if his wife were to say: “Look, it says butter and you have bought margarine”, he would hardly reply: “What a mistake! We must put that right, and alter the word on the list to “margarine”): whereas if the detective’s record and what the man actually buys do not agree then the mistake is in the record [Anscombe (1957), p. 56].

So, the shopping list, which has a “world-to-word” direction of fit, is analogous to a desire, which has a “world-to-mind” direction of fit. The detective’s inventory, which has a “mind-to-world” direction of fit, is analogous to a belief, which has a mind-to-world direction of fit. It is in the nature of belief to fit with how things stand in the world. If there is a disparity, it is the “fault” of the belief, so to speak. Conversely, it is in the nature of desire to motivate one to behave in such a way that will make the world fit with it. If there is a disparity, it is the “fault” of the world, so to speak.

Although this example is nice for illustrating direction of fit, I find Ramsey’s map metaphor more useful for understanding the essence of belief.<sup>6</sup> Note that the shopping list and the inventory both have the same format: both are lists, with discrete linguistic items. This suggests both a format for belief, and a similarity in format between belief and desire, neither of which I find plausible. Firstly, you have beliefs all the time, effortlessly, about everything you perceptually encounter. To put it another way, you are not (usually) informationally neutral about the things that you encounter. You are, however, *motivationally* neutral about the vast majority of what you encounter. We have appetites and goals. However, the vast majority of things (in the broadest sense to include objects and states of affairs) that we encounter aren’t the objects of those goals. Some of them might well be relevant to attaining some goal or other; others won’t even have *that* significance. Now consider the aptness of the map analogy. Suppose you want to get to point X, and realise that

you'll need to take such and such a route, through a tunnel, over a bridge etc. These are things that impact on you getting to point X. However, in looking at the map, you will also expect to see a church here, a forest there. These may be irrelevant to you getting to point X, but you still take them to be there.

Secondly, and in a related vein, beliefs are holistic, and necessarily subject to consistency in a way that desire doesn't seem to be. If I believe that you are shorter than 6ft, I will, by that same token, believe that you are shorter than 7ft, or 8ft etc. Maps, like beliefs, are holistic in this way. If there is a boundary between a forest and a field, and I update it on the map, making the forest bigger, I thereby make the field smaller. We do however seem to have conflicting desires all the time. I want to taste my friend's delicious chocolate brownies, but I know that they'll make me fat and I don't want to get fat. That's when decision-making gets tough. Granted, there are cases where believing in humans isn't maximally *consistent*. It is arguably, however, not possible for it to be directly *contradictory* (and in cases where it appears to be so, this can be explained away).

Furthermore, note that a map also has a map-to-world direction of fit. If there is a mismatch, it is the map's "fault". So the map analogy preserves the insight of direction of fit, without the infelicities of the list analogy that we have just mentioned.

### III.2. *Two Notions of Direction of Fit*

Direction of fit has been used in two slightly different ways, and it is important to disambiguate. According to one, direction of fit distinguishes the *informational* from the *motivational*, whereas according to the other, it distinguishes the *cognitive* from the *conative*. Sometimes these two oppositions are taken to be equivalent, but an illustration of the two distinctions will clarify the difference.

Velleman (1992) uses direction of fit in the second sense:

The term "direction of fit" refers to the two ways in which attitudes can relate propositions to the world. In *cognitive* attitudes, a proposition is grasped as patterned after the world; whereas in *conative* attitudes, the proposition is grasped as a pattern for the world to follow. The propositional object of desire is regarded not as fact – not, that is, as *factum*, having been brought about – but rather as *faciendum*, to be brought about: it's regarded not as true but as to be made true [Velleman (1992), p. 8].

The sense of "regarded" in "regarded as true", is a weak sense. As a result, Velleman's sense of "cognitive" includes any representational states that

present the world as being a certain way, *regardless of whether the subject actually takes it to be that way*. Thus, Velleman includes under the class of states or attitudes that are “cognitive”, and hence have mind-to-world direction of fit, states like perceptions, hypotheses, fantasies, imaginings etc. Indeed an attitude can have that direction of fit, and persist, even if the subject is totally aware of the mismatch between the state and the world (for example, when I’m imagining, or hallucinating, yet fully aware that I am).

The other use of “direction of fit” distinguishes, not the cognitive from the conative, but the informational from the motivational. This is the sense of direction of fit that John Searle has in mind. One very nice characterisation of it is by Michael Smith (1987). I cite it in full, and it rewards close reading.

...the difference between beliefs and desires in terms of direction of fit comes down to a difference between the counterfactual dependence of a belief and a desire that *p*, on a perception of *not p*: roughly, a belief that *p* is a state that tends to go out of existence in the presence of a perception that *not p*, whereas a desire that *p* is a state that tends to endure, disposing a subject to bring it about that *p*. Thus, we may say, attributions of beliefs and desires require that different *kinds* of counterfactuals are true of the subjects to whom they are attributed. We may say that that is what a difference in their direction of fit *is* [Smith (1987), p.54].

Such a use of direction of fit seems to *define* belief, *qua* basic informational state, *in terms of* direction of fit. Therefore, unlike Velleman’s use of “direction of fit”, states like imaginings, hypotheses etc. do not have a mind-to-world direction of fit.<sup>7</sup>

I will follow Searle, Smith, and others, in using direction of fit in this way, and in characterising belief in terms of the informational, understood as that which has a mind-to-world direction of fit in his sense. Another way of expressing the same thing as this sense of mind-to-world direction of fit is in terms of *transparency*. This is a term that gets used in various different ways, but the way in which I want to use it is as an expression of mind-to-world direction of fit. Beliefs are, by their nature, tied to (or simply are a function of) how the world is as far as the agent is concerned. As Gareth Evans puts it, belief is “directed outwards.” If you discover that *p* isn’t the case, you *ipso facto* cannot believe that *p*.<sup>8</sup> This is the same as direction of fit, since a belief that *p* will “go out of existence in the presence of a perception that *not p*”. Imagination that *p* is not transparent in this sense, and can certainly remain in existence “in the presence of a perception that *not p*”.

This has various revealing consequences for questions concerning whether you believe something. Let us contrast this with two other mental phenomena, for example, with desire (which has a world-to-mind direction of fit) or with an emotion (which *prima facie* has no direction of fit). If I ask you whether you are angry at the moment, you will need to assess yourself, for example, what you are feeling. However, if I ask you whether you believe that  $p$  (in the philosophical sense of “believe”) you don’t, or at least shouldn’t (unless you take me to be using “belief” in the everyday sense), look at yourself, as if trying to accurately describe an internal state (as you might with anger). Rather you simply ask yourself whether  $p$ .<sup>9</sup> This is obvious for perceptual beliefs, but it also extends to non-perceptual beliefs. If you ask me whether I believe that Paris is the capital of France, I ask myself whether Paris is the capital of France, not whether I believe it.<sup>10</sup> I examine the world, not myself. Belief is directed outwards, namely at the state of affairs that would make it true or accurate. These “outward” states of affairs can obviously include states of affairs involving one’s body, and one’s “self”, e.g. one’s personality.

So, belief is the basic informational state, believing is the basic informational phenomenon, and this involves taking the world to be a certain way. The question that interests us here is: What does one have to do (in the weakest sense of “do”) in order to count as taking the world to be a certain way? It is with this question in mind that I present my revisionist view about the norms of belief. For lack of a better name, I call it a “Downstream-only View”.

### III.3. *Evaluative and Constitutive Norms*

In order to answer the question of whether there can be belief without evidence, it is crucial to distinguish between evaluative and constitutive norms; in other words, distinguish between doing  $\phi$  badly or well (evaluative) versus doing  $\phi$  and not doing  $\phi$  at all (constitutive).<sup>11</sup> In this instance,  $\phi$ -ing is believing, and those who deny that there can be belief without evidence are claiming that this is breaking constitutive norms of belief. Something has gone so badly wrong (or wrong in such a way) that the subject is not longer believing at all.<sup>12</sup> In spite of her utterances, the subject doesn’t actually believe what she might seem to believe.

To get a grasp of the distinction between evaluative and constitutive norms, consider an analogy with chess (a real favourite among philosophers).<sup>13</sup> If, during a game of chess, I move so that my queen can be captured by my opponent with impunity (and as a result she gains an

enormous material advantage and goes on to win the game) that is a *bad* chess move. An evaluative norm (the norm of playing chess well) is broken. If I move my bishop like a rook, however, that's not a chess move at all, let alone a bad one. A constitutive norm (dictated by the rules of chess) is thereby broken. Those who want to deny the possibility of evidenceless belief think that it's like the latter. I think that it's like the former.<sup>14</sup>

### III.3. *Functional Roles and Norm-Breaking*

Let us consider what is typically seen as the functional role of belief. Beliefs are typically taken to:

- (i) Be supported by evidence in their initial formation,
- (ii) be open to review in the face of contrary evidence and
- (iii) guide action and reasoning.

Now, let us think of the functional role in terms of norms, and failure to fulfil those norms as being instances of breaking norms, namely, norms dictated by how a good belief ought (evaluative) or any belief must (constitutive) function. The norms (i)-(iii) divide into "upstream" and "downstream" norms. That beliefs are supported by evidence in their initial formation, and are responsive to review in the face of contrary evidence, constitute upstream norms, namely, norms that are related to causal or evidential precursors of the belief ((ii) may seem to be a downstream norm, since it is something that is exhibited after the belief has been formed, but since in attempting to correct the delusion, you are providing *input* for new judgement it is in fact upstream). That beliefs guide action or subsequent reasoning, i.e. (iii) is a downstream norm (namely, norms that govern the consequences of the subject's belief). Note that any belief-like state (like the religious commitment that "God is three and God is one") can be divided in this way. We can talk of how the subject came to be in such a state (upstream), and the use that the subject makes of the state (downstream).

Given what we have said about belief's transparent, map-like nature, what can we say about upstream and downstream considerations and their role in determining whether something is a belief? In other words, which of (i-iii) are constitutive norms, and which are evaluative? I suggest that there is a very simple pattern. Upstream norms are evaluative, whereas downstream norms are constitutive.

This claim, as we are about to see, is certainly controversial among philosophers, but it is ultimately correct and defensible. To reiterate: the claim is that, whereas upstream considerations are relevant for telling us whether the subject is believing well or badly, they are not relevant for telling us whether the subject is believing at all. In other words, to move back from talk of norms to talk of functional role, these upstream issues about sensitivity to evidence describe the functional role of *good* belief, not belief *tout court*.

This is, rather obviously, called the “Downstream-only View” because it is the view that a mental state (or propositional attitude) can be formed in any old way, but as long as it has the right kinds consequences (namely for action and dispositions for action) the state qualifies as a belief. Indeed the idea of a mad scientist manipulating your brain so as to “implant” a certain belief is not *prima facie* ruled out by our philosophical concept of belief.<sup>15</sup> This would be a prime case of irrational (or perhaps non-rational, depending on how you cash out epistemic irrationality) belief, but nonetheless it would count as belief, for example, if it were acted upon in the right way.<sup>16</sup>

Returning to the map metaphor, if we want to know whether the subject is using a map-like drawing as a map, we need to look at the subject’s dispositions to behave in accordance with the drawing. The presence of “disposition” is important, for, as we touched upon, there may be parts of the map, say distant parts, that are irrelevant to the subject’s goal-directed navigation, but the subject can still take them to be accurate, and this is shown by the fact that *if* the subject *were* to be in those distant parts, he *would* use the map to navigate (e.g. would expect a river here, a mountain there). If, for some reason, the map was drawn using a bad method (analogue: unjustified belief) or a good method but which unfortunately went awry and was consequently inaccurate (analogue: unluckily false belief), then the map is a *bad* map, but still used as a map if the subject is inclined to navigate by it. Recall that it is merely a metaphor: I am not saying that belief is a map, or even necessarily map-like in format. However, the notions of maps and navigation are useful ways to think about belief and the metaphor is particularly useful here, in the sense that the sort of questions that we ask for ascertaining whether something is being used as map, are the same sorts of questions that we ask to ascertain if something is a belief.<sup>17</sup>

#### III.4. *Two Objections to the Downstream Only View*

We have presented the downstream-only view. However, it is open to two compelling, but ultimately misguided, objections. The first objec-

tion is an objection from strong normativism. This objection is essentially a straightforward contradiction of the downstream-only view, which is grounded in the intuition that belief “aims at truth”. The second objection is the objection from alternative explanations. Belief can’t simply be about dispositions to actions, since there are other states that intuitively aren’t beliefs but that explain action in the same way. Hence we need upstream considerations to tell these belief and non-belief states apart. I will present and respond to these two objections in turn in sections IV and V respectively.

#### IV. OBJECTION 1: STRONG NORMATIVISM AND AIMING AT TRUTH

The objection from strong normativism involves the claim that belief, by its nature, “aims at truth”, or is constitutively subject to the “truth norm”. We can break up this objection into three steps:

1. Beliefs, by their nature, aim at truth.
2. Aiming at truth is an upstream consideration (failing to aim at the truth infringes an upstream norm).
3. Therefore the downstream-only view is mistaken

I will state my response to this first, and then flesh it out. It is as follows.

1 is ambiguous, in particular with regards to the notion of “aiming at truth”. In one sense of “aiming at truth”, 1 is true. However, in the sense in which 1 is true, 2 is false. Conversely, in the sense in which 2 is true, 1 is false. Namely, once “aiming at truth” is conceived as an upstream consideration, it ceases to be constitutive of belief. I will now flesh this out by looking at some interpretations of 1.

##### IV.1. *Believers “aim at truth” in Believing*

So, in what sense can we mean that belief, by its nature, aims at truth? Well, let’s start by highlighting what *cannot* be meant by this.

What cannot be meant by this is that the believer, in believing, literally, directly, tries to have beliefs that are true. This can’t be true because belief is involuntary. The believer can’t aim at *anything* in believing, or at least not *directly*. If one is not convinced that belief is involuntary (although in my view, one should), there is another argument against this reading of the truth-aim claim, which doesn’t rely on involuntarism (but which does provide indirect support for it). If the truth-aim claim is seen

as an aim that can actually guide the subject's beliefs then it suffers from the following problem (due to Gluer and Wikforss (2010)). Consider other norms or aims that guide behaviour. For example, consider a rule that a stock-market trader might use: "buy low and sell high". In order for this to actually guide the subject, the subject needs a belief about the current state of the stock market (namely, that what one is considering buying or selling is low or high). Now consider the norm: "Only believe something if it is true". In order for this to guide belief formation, you already need to have a belief about whether the thing in question is true. But then you already have a belief about the matter in question. Therefore the norm "comes too late" to guide belief. Needless to say, this falls out of transparency. And involuntarism also falls, in turn, naturally out of this, since belief is not an action, not something that the agent does (and hence not something the agent can be guided in doing), but rather a *precondition* for action.

#### IV.2. Believers "aim at truth" by Regulating for Truth

Perhaps what is meant is something more *regulative*. So, to think of an analogy, I can't directly control my cholesterol, but I can *aim* at lowering my cholesterol indirectly by observing a cholesterol-lowering diet. Similarly, with belief, I can't directly try to believe things that are true, but I can aim at truth indirectly by observing good epistemic practice, for example, by not interpreting or gathering evidence in a biased way etc. However, although this is now a coherent way of understanding the truth-aim claim as something that the *subject* aims for, namely, she aims for it in a regulative way, we now plausibly lose the *constitutive* force of the claim. And the claim is supposed to tell us what is constitutive of belief. What we now have is rather an evaluative, rather than constitutive norm. Biased evidence gathering, failure to "regulate" for truth, leads to *bad* believing, but not to failure to believe altogether.<sup>18</sup>

What we might say, however, is that it is *constitutive* of belief to have certain *evaluative* norms. This seems plausible. But then it is not constitutive that the believer aim at truth, but rather that she *ought* to. The consequences that this has for a belief-like state that is the product of biased reasoning, or that is poorly grounded, or unresponsive to evidence, is that this is believing that is *bad*, that the subject ought not to have, but it is still belief. This is perfectly consistent with a downstream-only view.

#### IV.3 Belief qua mental state "aims at truth"

One last interpretation, which follows on from what we have just said, is that belief *itself*, the mental state, in some metaphorical sense *aims*

*at truth*. This avoids problems about involuntarism, and about the norm coming “too late”. So, what is meant by this metaphor? We can see it all the more clearly when we compare it to other cognitive mental states (by “cognitive” I mean, as before, as opposed to “conative”: the world is represented as being a certain way rather than to be made that way).

Consider a contrasting cognitive state of choice, viz. imagining. An imagining that misrepresents the world, or is not grounded in evidence, is not “bad” in any way. One can say that this is because imagining (the state itself) is under no “obligation” to “aim at truth”.<sup>19</sup> What this seems to suggest is that it is not the believer who aims at truth but in some metaphorical sense, the *state itself*, which by its very nature “aims at truth” (and that in turn constitutively entails an evaluative norm on the part of the believer). This falls out of the state’s direction-of-fit, understood in the transparent, informational sense. If this claim is true, it is trivial and perfectly in keeping with a properly understood downstream-only view. Belief aims at truth insofar as it is how the subject takes the world to be, but it still doesn’t strictly matter for belief-status how the subject came to take the world to be that way.

One final ambiguity that needs ironing out, which is present, both in talk of “transparency” and “aiming at the truth”, is whether we are talking about objective, *de facto* transparency or truth-aiming, or merely *as far as the subject is concerned*. I think it is fairly obvious that we mean the latter.<sup>20</sup> If we meant the former, we would have to deny belief-status to mistakes (faultless, unlucky mistakes, as well as biases). And yet, mistakes get all their import from being doxastic, from being beliefs (which amounts to them being transparent). A mistake, a misrepresentation of reality in imagination, is not problematic because it isn’t purporting to represent reality, and so typically wouldn’t be acted upon in the service of relevant goals. So, to clarify, when we say that belief aims at truth, we mean truth as far as the subject is concerned. When we say that it is transparent, directed at the world, at reality, we mean the world, reality, as far as the subject is concerned. To use Ramsey’s metaphor, we want to know what the subject’s map is like, and an inaccurate map, even a systematically inaccurate one, is still a map.

#### IV.4. Recap of the Response to Objection 1

As I mentioned, “aiming at the truth” is ambiguous. If this means that the agent aims at the truth, then either this makes no sense, if the aiming is supposed to happen directly, or it does make sense and is regulative. If it is regulative, it fails to be constitutive of belief. Another inter-

pretation of “aiming at truth” is that the state itself aims at truth. This is merely a consequence of its direction of fit, and is perfectly in keeping with the downstream-only view.

## V. OBJECTION 2: NON-BELIEF STATES EXPLAIN ACTIONS

This objection runs as follows. People often act in ways that are best explained in terms of states that aren’t beliefs. How do we distinguish these states from beliefs? The answer is: by using upstream considerations, such as their relationship to evidence etc. Therefore the downstream-only view is false.

What are these non-belief states? We will divide them into two groups. In one group, let’s place those that are referred to using pre-existing words in English. In another let’s place those that are referred to using terms of art. I will examine these in order, and then respond to both together.

### V.1. *Imagining, Pretending and Believing*

The criticism goes as follows. There has to be more to belief than mere dispositions to action, since there are states that fall short of belief, like imaginings and pretendings, but which have similar behavioural consequences. Thus, when a child is having a pretend tea party, we don’t attribute the belief that she’s pouring tea, even though the behaviour is similar. We say that she is imagining that there is tea in the teapot, and engaged in pretence. We say similar things about actors on stage. We might say that what makes these cases of imagining and not belief is the way that they are related to evidence. The child is not believing that there is tea in the teapot since she can see that there isn’t. If you were to try to correct the child’s behaviour by pointing out that there is no tea in the teapot, she’d quite rightly be somewhat confused by your attempt to correct her. This is because imagination and pretence are not responsive to evidence. One might think that beliefs, contrary to this, are thus responsive. If the child did believe that it was tea, and you pointed out that it was merely water, she might taste it, realise that you are right, and revise her beliefs. But the child is not believing. She is imagining or pretending.

### V.2. *Alief and Bimagining*

I’ll introduce this with the opening example from Tamar Gendler’s well-known paper. In March 2007, a glass walkway, suspended 4000 feet

above the Grand Canyon, was opened to the public. Most tourists struggle to make it to the centre of the walkway; others are paralyzed with vertigo.

How should we describe the cognitive state of those who manage to stride to the Skywalk's center? Surely they *believe* that the walkway will hold: no one would willingly step onto a mile-high platform if they had even a scintilla of doubt concerning its stability. But alongside that belief there is something else going on. Although the venturesome souls wholeheartedly *believe* that the walkway is completely safe, they also *alieve* something very different. The alief has roughly the following content: "Really high up, long long way down. Not a safe place to be! Get off!!" [Gendler (2008a), p. 635]

Gendler has used this and several other examples to motivate putting forward the well-received view that there are belief-like states, that are not beliefs, but which have to be included into our taxonomy of mental states for explanatory reasons (of the same kind that motivate the presence of belief in such a taxonomy). Roughly, to have an *alief* is to have "an innate or habitual propensity for a real or apparent stimulus to automatically activate a particular affective and behavioural repertoire, where the behavioral propensities to which an alief gives rise may be in tension with those that arise from one's beliefs" [Gendler (2008b), p. 553]. Andy Egan (2009) posits a similar hybrid (or "in-between") state, "bimagination", to explain roughly similar phenomena (including delusions). Bimagination has some of the functional role of belief and some of the functional role of imagination: it is both action-guiding (and hence belief-like) but resistant also to counter-evidence (and hence imagination-like.)

There has been some dispute about whether, and to what extent, *alief* is explanatorily necessary or helpful. In particular, critics, [see e.g. Mandelbaum (2013)] think that the relevant phenomena can be explained by beliefs and desires coupled by interference from subpersonal processes (e.g. cued responses, conditioned or innate, to certain stimuli). Although I think my preferred view of belief (and its role in explanation) may be conducive to supporting such criticisms, such criticisms are not strictly relevant to defending the downstream-only view. The point is, rather, that Gendler takes these cases, and her solution to these cases by using aliefs, to oppose, and cause serious problems for something she calls a "Neo-behaviorist view" [Gendler (2008a), p. 653]. Gendler characterizes such a view as follows:

... to believe that *P* is to be disposed to act in ways that would tend to satisfy one's desires, whatever they are, in a world in which *P* (together with one's other beliefs) were true [Gendler (2008a), p. 652].

She says that the real problem for this view is that fact that “a wide range of attitudes – among them acceptance (Bratman), imagination (Currie, Velleman) and pretense (Doggett & Egan, Velleman) – may motivate P-concordant behavior” [Ibid., p. 653]. She wants to add herself and alief alongside these philosophers and attitudes just cited. What distinguishes these states from being belief that P cannot simply be the way that they motivate behaviour, since, by stipulation, they motivate the same behaviour (namely “P-concordant” behaviour). What makes these states differ from beliefs is that belief has a certain relationship to evidence (it is, as Gendler puts it, *reality-sensitive*) that these other states lack. At this point it should be clear that accepting that belief requires a certain relationship to evidence, and not simply action, amounts to giving up the downstream-only view of belief. And, of course, that is the view that I am trying to defend.

I will respond to the objections from imagination and from alief in reverse order, so, starting with alief. The same central principle guides my responses to both, namely, an important distinction between action and behaviour.

### V.3. Responses to Objection 2: Action and Mere Behaviour

Let Gendler have aliefs, and let’s suppose that they are worth positing. What is my response to her critique of “Neo-behaviorism”? Firstly, a note on “Neo-behaviourism.” If by “Neo-behaviourism” she means simply “behaviourism that happens to be held today”, then I would not call the downstream-only view, “behaviourism”. And yet, I entirely agree that her characterization of “Neo-behaviourism” (see quotation above) is the downstream-only view. So what’s going on?

Behaviourism, as Chisholm (1957) rightly pointed out, is a reductive project. It aims to rephrase, *salva veritate*, claims about the mental in terms of claims about the non-mental, and, in particular, the behavioural. Thus, claims about belief are rephrased (“analysed”, if you will) in terms dispositions to behaviour. However, the downstream-only view is not a reductive enterprise in this sense at all. We are not concerned here, for example, with a metaphysical project of trying to accommodate the mental in a physical world. We are simply trying to state the necessary conditions for someone to count as a believer. Thus we rephrase claims about believing in terms of claims about *action*. But action is a mental phenomenon. We are not being reductive. We are elucidating one mental phenomenon in terms of another mental phenomenon. We are picking out necessary relations between them. Well, what is the purpose of this here? So that we can tell when someone appears to be believing something when in fact they are

not. To take an analogy, saying that, by knowing what someone desires and how they are disposed to act, you can ascertain what they believe, is like saying that you can ascertain velocity from knowing distance covered and time taken. That is not a *reduction*. It does not “reduce” velocity to distance and time any more that one can “reduce” distance to velocity and time. You may have views about which is metaphysically more fundamental, but the relations between the phenomena will hold independently of this. Similarly, you may think that action has metaphysical primacy, or that belief has, or desire has. But that is a different issue.

Now note that Gendler’s characterization of “Neo-behaviourism” isn’t reductive *either*. In it, she talks of disposition to *action*. But then she goes on to talk about P-concordant *behaviour*. She slips from talk of action to talk of behaviour, without marking the difference. And yet there is an enormous difference. If she had characterized neo-behaviourism in terms of behaviour it would be reductive and, I agree, erroneous. However, *action* is, at least in part, individuated by intention, not by mere bodily motion. Although there has been a great deal of debate surrounding the precise individuation conditions of actions, here we only strictly need to note something that all views agree on: that the same bodily motions performed with different intentions are different actions. Behaviour, on the other hand, is individuated by bodily motion. A result of this is that two identical pieces of outward behaviour can be different actions. They will be different actions if the intentions are different, and the intentions will be different if the beliefs or desires are different. We can contrast cases where intentions are lacking, or when they are different.

Suppose, in Scenario 1, John kills (in a neutral sense of being causally implicated in his death) James, because Joe, who wanted to frame him, implanted a microchip in his brain and controlled his behaviour. In Scenario 2, John kills James as a way of exacting revenge. We can stipulate that the outward behaviour is the same in both scenarios. However, the two scenarios are very different. In Scenario 1, we wouldn’t think of John’s causing of James’s death as an action, at least not John’s action. We might even plausibly call it Joe’s action. This would be reflected in our moral and legal practices. Given all the relevant information, we would blame Joe for James’s death and charge him with murder. In Scenario 2, we simply have the straightforward action of John murdering James. So, actions are different actions (sometimes, as we saw, attributable to different *individuals*) if the intentions are different. Bodily movement, without any intention, fails to be action at all.

One query, worth addressing now, however, is: how do we make space for the notion of unintentional action? This doesn't seem like a contradiction in terms, and yet it seems that on the view just sketched, this is ruled out. In order to respond to this, it is important to distinguish the action, from the way the action is described. As Davidson (1980) influentially put it, in order for something to be an action, it has to be "intentional under some description". Thus, to take an example from Searle (1983), it seems intuitively plausible that Oedipus's marrying Jocasta is an action. An action performed with the intention of marrying a beautiful woman whom he loved. It was not performed with the intention of marrying his mother, although that is *de facto* what he did. So we might say that he unintentionally married his mother. However, note that it would fail to be an action at all if there was *no* description under which it was intentional. Due to our ignorance, we often perform actions successfully, but with unintended consequences.

So here is my diagnosis of what aliefs show for the downstream-only view. Gendler is right to say that these states are not beliefs. However, this is not, as she says, because they are not supported by evidence or inflexible in the face of contrary evidence. There are, as I argue (and as Bortolotti (2009) was at pains to point out) many beliefs that are resistant to correction (although they should only be thought of as beliefs if the subject is disposed to act on the in the right way). Aliefs fail to be beliefs – or rather, subjects in the examples do not believe the "content of the alief" – because these states *causally explain behaviour* rather than *rationaly explain action*. Aliefs are affective states that *interfere* with people's intentions and actions, rather than taking part in them in the way that beliefs typically do. The person reluctant to walk on the walkway, wants to walk on the walkway, believes that it is safe, but is hindered by a strong sense of fear. The person reluctant to drink from the bedpan [Gendler (2008a), p. 636], may want to drink from the bedpan (presumably the experimenter gives them some incentive to do so), believes that it is safe to do so, but is hindered by a strong feeling of disgust.

Now let's look at imagination. As we saw, Gendler's objection to "Neo-behaviorism" is confusing the individuation conditions of action and mere behaviour. The same response applies here. Part of what makes something imagination or make-believe, rather than belief, is not, as Currie emphasizes, that it's not responsive to epistemic reasons, but rather that the action on the basis of which you ascribe the merely imaginative episode in question is different from the action on the basis of which you would ascribe a state of believing (even if the *behaviour* is su-

perfidially the same). An actor on stage who runs away from something dangerous in the play doesn't believe that he is in danger, but rather believes that he is acting in a play, wants to give the audience the impression of being in danger, and believes that this is how that's done. He therefore isn't literally performing the action of running away from danger, however convincing he may be in his acting. Performing the action of running away from danger requires that one have the belief that one is in danger. Otherwise one is not performing *that* action, but merely something that superficially resembles it.

## VI. CONCLUSION

In this paper, I presented a downstream-only view of belief, according to which there can, indeed, be belief without evidence. This neither depended on there actually being cases of belief without evidence (although I'm sure there are such cases), namely, it was a mainly conceptual rather than empirical exploration. Nor was it dependent on our intuitions about whether we'd be happy to attribute belief in such cases, namely, the view is revisionist, not conservative. For example, the view would lead to revisions in cases where people make sincere assertions, but do not act in accordance with those assertions. Although I haven't the space to elaborate on these example, this could be due to a lack of self-knowledge, lack of linguistic knowledge, or lack of understanding. Several of our so-called "beliefs" fall short of actually being beliefs, but that is because we do not, and sometimes cannot, act upon them. It is not because they lack evidence.

*Department of Philosophy  
Durham University  
50 Old Elvet, Durham DH1 3HN, UK  
E-mail: sam.wilkinson@durham.ac.uk*

## NOTES

\* This paper was a finalist entry at the 2014 **teorema** Essay Prize for Young Scholars on "Belief Without Evidence".

<sup>1</sup> She is open to thinking about propositions in terms of possible worlds. But although her view has all the same consequences of such a view, she claims that she is more comfortable with the vocabulary of states of affairs.

<sup>2</sup> Note, also, the potential implications for philosophy of mathematics: when a mathematician makes a mistake, we intuitively say that they used to believe something that was impossible (namely a mathematical falsehood). This is ruled out by such a view.

<sup>3</sup> Whether informational states in an artificial system, i.e. a robot, can be thought of as beliefs is an interesting question, but one that is not strictly relevant for our purposes. I would suggest that good reasons for opposition to this view would not be something intrinsic to the informational states themselves, but rather because the goals are not theirs. They are the goals of the programmer. Artificial systems only have derived goals, namely, the goals of their creators. Evolved beings have goals of their own.

<sup>4</sup> What Ramsey meant by “we” is not entirely obvious. He could have meant “human beings”. I’d like to think (and it fits in with his view, and the apparent modal strength of his claim) that he meant “any beings capable of belief”.

<sup>5</sup> Austin used it as a way of explaining different ways of asserting  $Fa$  (and given this difference, the logical form is insufficient to capture it). This is seen most clearly when we assert something of the logical form  $Fa$  which is false. Compare wrongly calling a triangle a square, which is “committing an act of violence to the language”, with wrongly describing a triangular object as being square, which is “committing an act of violence to the facts”. This difference is explained in terms of direction of fit since in the former the fault is in “fitting a name to the item”, whereas in the latter it is in “fitting the item to the name”.

<sup>6</sup> Beliefs with modal or normative content may be hard to capture in literal maps, for example. However, they only qualify as beliefs if they guide my actions, and *in this respect* are map-like. If I truly believe that it might rain, I will, *ceteris paribus*, bring my umbrella. If I truly believe that killing babies is very wrong, I will, *ceteris paribus*, refrain from doing it. Perhaps, in any case, we can make sense of maps with normative and modal content. A street on a map with an arrow indicating one-way traffic has normative content. One with a warning sign that flooding at high tide is a risk has modal content (of sorts).

<sup>7</sup> Indeed it is not clear whether in this sense these states have an essential direction of fit at all - they may have elements of a mind-to-world direction of fit, to the extent that they have elements of *belief*, to the extent that they are *used as information*.

<sup>8</sup> See also Moore’s famous paradox. Although the point is very similar, however, it is phrased in semantic rather than epistemic terms.

<sup>9</sup> Velleman uses “transparency” to refer to a different phenomenon. He claims that the question “Whether to believe that  $p$ ” is, “transparent to”, namely, answered by answering the question “Whether  $p$ ”. We might call this *deliberative transparency*. The transparency I am talking about here is what we might call *transparency of self-attribution*. The latter tells us something more fundamental about the nature of belief. Furthermore, although I don’t need to get embroiled in this tricky debate, I don’t agree with *deliberative transparency* since I don’t believe in

*doxastic deliberation*, namely, we never ask whether to believe that *p*, since belief is involuntary in the strongest sense (due to its nature, and in all possible worlds).

<sup>10</sup> Although in everyday speech, people talk about “beliefs” as the inconclusively supported commitments that help to define them as a person: e.g. “I stand by my beliefs”. This is clearly not what I mean. People may well be asking questions about themselves when asked what they “believe” in these cases.

<sup>11</sup> I mean “doing” in the weakest sense. It does not imply action or intentionally doing something.

<sup>12</sup> The parenthesis “or wrong in such a way” is to highlight that one might not think of this as a matter of degree. Some theorists, in a way analogous to asking “How irrational must one be in order to count as delusional?” may ask: “How irrational must one be in order for one to fail to be a believer?” Others, rightly in my opinion, will see the believing/not-believing boundary as a difference in type rather than degree.

<sup>13</sup> Wedgwood (2002), for example, uses the chess analogy.

<sup>14</sup> This analogy with chess, however, has limited utility and it is important to locate where the analogy breaks down. Belief is like chess in that it is subject to norms; some of the norms will be essential to what belief is, others will feature in evaluations of beliefs. In both cases, the constitutive norms being adhered to are a precondition of the evaluative norms coming into play. A move obviously can’t be a bad chess move if it’s not a chess move at all. And the same goes for belief. But the parallel stops there. Disanalogies, for example, include the fact that chess is a game, the playing of which is utterly deliberate. A move in chess is an *action*, rendered intelligible by beliefs and goals (by informational and motivational states). Given this fact about chess, you can deliberately play badly if your goal is to do so. As we will see, it is not obvious that this can be said of belief, at least not directly.

<sup>15</sup> As we will see in the next chapter, for me, this would be not the implanting of a discrete sentence-like entity, but the “implanting” (or, better, instilling) of a particular disposition to action.

<sup>16</sup> Some would say that you needn’t go as far as farfetched thought experiments: the brain-washing of extremist cults, for example, is behaviour-driving in the worst ways (e.g. suicide bombing), but the belief-formation process is far from rational.

<sup>17</sup> There is an illustrative limit to the map/belief analogy. Maps are typically things that you deliberately *use* to help you navigate. Beliefs are not things that you *use*: they are things that you *have* when you are in the process of navigating. Thus, it is more helpful to think of maps in the analogy *as models* rather than *as literal maps*. In other words, the subject isn’t going through a process of using something map-like to navigate. Rather, the agent’s informational state, as manifested in her goal-directed navigations, can be *modelled* (roughly, helpfully described) in terms of something map-like.

<sup>18</sup> This regulative truth aiming may, however, be a constitutive norm for a more active evidence-gathering process, like “inquiry”. In a strong, but not out-

landish sense of “inquiry”, somebody who treats evidence in a biased way may not be inquiring badly, but failing to inquire at all. Is someone who plays chess, put tries to lose, playing chess or not? Well, yes and no. This is not a substantive issue, but a terminological one: do you include the attempt to win in the concept of “playing chess”? Does it matter either way?

<sup>19</sup> As Velleman (2000) puts it, unlike imagining, “belief is the state that goes right or wrong by being true or false”.

<sup>20</sup> And arguably this is captured in the metaphor of “aiming”.

#### REFERENCES

- ANSCOMBE, G. E. M. (1957), *Intention*, Oxford, Basil Blackwell.
- AUSTIN, J. L. (1962), *How to do Things with Words*. The William James Lectures delivered at Harvard University in 1955, (ed. J. O. Urmson and Marina Sbisa); Oxford, Clarendon Press
- BORTOLOTTI, L. (2009), *Delusions and Other Irrational Beliefs*, Oxford, Oxford University Press.
- CHISHOLM, R. M. (1957), *Perceiving*; Ithaca, Cornell University Press.
- CURRIE, G. (2000), “Imagination, Delusion and Hallucinations,” in M. Coltheart and M. Davies (eds.) *Pathologies of Belief*; Oxford, Blackwell, pp. 167-182.
- DAVIDSON, D. (1980), *Essays on Actions and Events*; Oxford University Press.
- EGAN, A. (2009), “Imagination, Delusion and Self-Deception” in T. Bayne and J. Fernández (eds.), *Delusion and Self-deception: Affective and Motivational Influences on Belief Formation*; Hove, Psychology Press.
- GENDLER, T. (2008a), “Alief and Belief,” *Journal of Philosophy*, 105(10), pp. 634-663.
- (2008b), “Alief in Action (and Reaction),” *Mind and Language*, 23(5), pp. 552-85.
- GLUER, K. & WIKFORSS, A. (2010), “The Truth Norm and Guidance: A Reply to Steglich-Petersen”; *Mind* 119 (475), pp. 757-761.
- MANDELBAUM, E. (2013), “Against Alief?,” *Philosophical Studies* 165 (1), pp. 197-211.
- MARCUS R. B. (1995), “The Anti-Naturalism of Some Language Centered Accounts of Belief?,” *Dialectica* 49, pp. 113-130.
- RAMSEY, F. P. (1931), *Foundations: Essays in Philosophy, Logic, Mathematics, and Economics*; Humanities Press.
- SEARLE, J (1979). “What is an Intentional State?” *Mind* 88 (January), pp. 74-92.
- (1983). *Intentionality*; Oxford, Oxford University Press.
- SMITH, M. (1987). “The Humean Theory of Motivation”; *Mind* 96 (381), pp. 36-61.
- VELLEMAN, J. D. (1992), “What Happens When Someone Acts?” *Mind* 101 (403), pp. 461 - 481.
- (2000). “On the Aim of Belief?,” In David Velleman (ed.), *The Possibility of Practical Reason*; Oxford, Oxford University Press.
- WEDGWOOD, R. (2002), “The Aim of Belief?,” *Philosophical Perspectives* 16, pp. 267-297.