Aiming at Truth: On The Role of Belief

Kathrin Glüer and Åsa Wikforss

RESUMEN

En este artículo exploramos la posibilidad de caracterizar la creencia en términos totalmente de su papel funcional de primer orden, su *input* (evidencia) y su *output* (creencias y acciones subsiguientes), al abordar algunos desafíos comunes a este punto de vista. Uno de los desafíos tiene que ver con el hecho de que no toda creencia es sensible a la evidencia. Como respuesta a esto, los normativistas y los teleo-funcionalistas han concluido que se necesita algo además del rol funcional, una norma o un *telos*. Argumentamos que ambas cosas permiten de forma implausible demasiada divergencia entre creencia y evidencia. Otros han sugerido que la creencia debería salvarse como *la* actitud sensible a la evidencia, haciendo que comparta su papel motivacional con un estado hasta ahora no reconocido: la acreencia. Argumentamos que la apelación a la acreencia tiene que hacer frente a un dilema. O la explicación de la acción intencional por medio de acreencia es una especie de explicación intencional, en cuyo caso resulta difícil distinguir acreencia de creencia (irracional), o la acreencia es suficientemente diferente de la creencia, pero entonces ni la explicación ni el *explanandum* (acción) son ya reconociblemente intencionales. Concluimos que el modo más prometedor de avanzar es una explicación de la creencia que haga uso del completo papel funcional de la creencia, incluyendo su papel en el razonamiento teórico.

PALABRAS CLAVE: *creencia, normativismo, teleo-funcionalismo, razones, explicación de la acción, acreencia*.

ABSTRACT

We explore the possibility of characterizing belief wholly in terms of its first-order functional role, its input (evidence) and output (further beliefs and actions), by addressing some common challenges to the view. One challenge concerns the fact that not all belief is evidence-sensitive. In response to this, normativists and teleo-functionalists have concluded that something over and above functional role is needed, a norm or a *telos*. We argue that both allow for implausibly much divergence between belief and evidence. Others have suggested that belief should be saved as *the* evidence-sensitive attitude, by making it share its motivational role with an hitherto unrecognized state: alief. We argue that the appeal to alief faces a dilemma: Either explanation of intentional action by means of alief is a species of intentional explanation, in which case it becomes hard to distinguish alief from (irrational) belief, or alief

is sufficiently different from belief, but then neither the explanation nor the explanandum (action) are recognizably intentional any longer. We conclude that the most promising way forward is an account of belief that makes use of the full functional role of belief, including its role in theoretical reasoning.

KEY WORDS: *Belief, Normativism, Teleo-Functionalism, Reasons, Action Explanation, Alief.*

I. INTRODUCTION

Belief aims at truth. Ever since Bernhard Williams first used this poetic metaphor – with its evocation of a little *Amor Veritatis* lifting his bow – it has cast its spell over the philosophy of mind. As befits a metaphor, it has been interpreted in various, not always mutually compatible ways. And even though the upshot of this paper will be largely negative, we, too, are very fond of the metaphor. Belief indeed aims at truth. There is nothing very lofty about the literal truth behind the metaphor, however: Belief aims at truth by playing a distinctive first-order functional role.

This role can be abstracted from a coherent and sufficiently comprehensive theoretical reconstruction of the use made of the notion of belief in folk-psychological generalizations, explanations, and predictions – an idea familiar from the writings of David Lewis and Donald Davidson. It will have two main components: What can be called belief's "input-" or "formation-role", and its "output-" or "dispositional-role". Part of the latter will be belief's motivational role, i.e. the role it plays in action motivation and explanation. Very roughly, belief aims at truth in that its input-role is characterized by a distinctive *evidence-sensitivity*, and its output-role such that it provides a distinctive sort of *reasons* both for further belief and for action.

But this is not the place either to formulate or defend a precise account of belief in terms of its functional role.[1] In previous papers (2009, 2010, 2013), we have argued that it is not a good idea to construe belief as essentially normative, where 'normative' is construed in terms of prescriptions guiding belief formation. This, we claim, is neither necessary nor feasible, thus clearing part of the way for a descriptive, functionalist account of belief. In this paper, we shall be concerned with more such roadwork.

A number of writers have argued recently that the traditional picture of belief's motivational role must be radically revised. The traditional picture – as pioneered by Davidson in a series of seminal articles beginning with his (1963) – forges very tight connections between belief and intentional action: Belief's motivational role is to 'direct' desire to its satisfaction. Together, belief and desire thus provide intentional explanations for actions. These explanations are reasons-explanations, and according to Davidson, *every* intentional action has a reasons-explanation [cf. Davidson (1971)].

Recently, this picture has come under sustained pressure. It has been argued that belief cannot be characterized in terms of its motivational role because there are a number of other mental states that have the same motivational role: assumption, imagination, and alief. Its motivational role thus is taken to be necessary for belief, but not sufficient. Why not combine motivational role with formation role, then? Assuming, imagining, and alief certainly are not evidence-sensitive in the sense in which belief typically is.

At this point, the line of argument forks. Those following the more well-travelled branch will respond that while having a certain motivational role is not sufficient for being a belief, being evidence-sensitive is not *necessary*. Not all belief is formed on the basis of sufficient evidence, and some belief, it is often held, might even be completely evidence-immune. To be sure, belief *aims* at truth, but that cannot mean that each and every belief is formed on the basis of sufficient evidence, or even on the basis of any evidence whatsoever. Consequently, trying to characterize belief in terms of its evidence-sensitivity is bound to get the extension of belief wrong. We shall call this "the extension problem". To solve the extension problem, this line of argument concludes, something over and above functional role is needed, something such as a norm or a *telos*.

At a certain level of abstraction, both normativism and teleofunctionalism about belief are attempts to reconcile the idea that evidence-sensitivity is essential to belief with the fact that actual belief formation does not (fully) conform to this ideal. In the second and third part of this paper, we shall argue that the gap between ideal and first-order reality that teleofunctionalism introduces becomes too big. What is intended as a little distance to keep the relationship intact ends up in divorce. We shall also offer a quick and dirty application of the very same diagnosis to normativist accounts of belief. This speaks in favor of having another go at trying to make the idea that belief can be characterized in terms of both parts of its first-order functional role work.

The other, less well-explored response to this idea can in fact be construed as precisely such an attempt. In the fourth part of the paper, we shall thus look at an intriguing new proposal as to how to 'save' the traditional philosophical notion of belief as *the* evidence-sensitive propositional attitude by making it share its motivational role with a hitherto unrecognized, largely evidence-immune mental state: that of alief. "Aliefism" tries to solve the extension problem by reserving the notion of belief for precisely those states that do show the right kind of evidence-sensitivity. But even though aliefism is wonderfully first-order, we worry that it cannot save traditional belief. Rather, it faces a dilemma.

Aliefs are supposed to be like beliefs in motivating and explaining what is not 'just behavior' but intentional action – at least in a wide sense of intentional action. But then, the traditional notion of intentional action needs to be quite radically modified: On the new picture, having a belief-desire explana-

tion will no longer be necessary for being an (intentional) action. For the new picture to be plausible, however, action explanation by means of alief ("alief explanation") needs to be sufficiently similar to belief-desire explanation in order to even be recognizable as a form of intentional explanation (in a wide sense), and for its explanandum to be recognizable as intentional action (in a wide sense). At the same time, alief needs to be sufficiently different from belief to be recognizable as a different kind of mental state. But, or so we shall argue, these two desiderata cannot be met at the same time. That's what we shall call "the alief-dilemma": Either alief-explanation is recognizably intentional. Then alief becomes too much like (irrational) belief. Or alief is sufficiently different from (irrational) belief. Then alief-explanation is no longer recognizably intentional.

It therefore seems to us that all the main ways of construing belief's aiming at the truth currently on the market lead into dead ends. But why should we think that this clears the way for a more traditional functionalism about belief? Rather than conclude that functionalism itself is at a dead end? We shall end the paper by observing that none of the positions investigated makes use of the *full* functional role of belief. On the input side, evidence-sensitivity is stressed, and on the output side motivation. What is neglected is the role of belief in theoretical, as opposed to practical reasoning, especially its output side. But belief has a characteristically truth-directed reason-providing role not only for action, but also for further belief. Therefore, the hope for characterizing belief in terms of its distinctive evidence sensitive, reason providing role for both action and further belief is not out yet.

II. TELEOFUNCTIONALISM AND THE AIM OF BELIEF

In his well-known discussion of whether it is possible to believe at will, Bernard Williams characterizes the nature of belief in terms of five central features [Williams (1973), pp. 136-151]. The first feature, Williams suggests, can be roughly summarized by saying that belief 'aims at truth'. The metaphor is employed by Williams to capture various aspects of the close connection between belief and truth. For instance, truth and falsehood are a dimension of assessment of belief, as opposed to many other psychological states. Moreover, to believe that p is to believe that p is true. If a man recognizes that what he has been believing is false, Williams says, "he thereby abandons the belief he had" [Ibid. p. 137]. Williams then argues that this feature of belief, its aiming at truth, is one central reason why we cannot believe at will. If one could believe at will then one could adopt the attitude of belief towards p regardless of the truth of p and this is incompatible with the fact that beliefs aim at truth [Ibid. p. 148].

David Velleman's teleological account of belief takes Williams' metaphor as its starting point [Velleman (2000)]. Like Williams, Velleman suggests that belief can be characterized in terms of its aiming at truth, but he suggests that the metaphor of aiming is to be taken more literally, as involving a *purpose* or *goal* that is distinctive of belief. The mere fact that there is a connection between believing and believing something to be true, Velleman argues, does not capture the truth-directedness of belief, since this is something belief has in common with a large number of other attitudes, such as wishing, desiring and hoping something to be true. Nor does it help adding that belief (unlike the conative attitudes) involves *regarding* something as true, Velleman argues, since regarding something as true is involved in the other cognitive attitudes as well; for instance supposing, assuming and imagining that *p* all involve regarding *p* as true, or the attitude of *acceptance* [Ibid., pp. 249-250]. This attitude of acceptance, Velleman suggests, can be defined in terms of its motivational role. To accept *p* is to have a disposition "to behave as would be desirable if the proposition *were* true, by doing things that would promote the satisfaction of one's desires in that case" [Ibid., p. 255]. This is the motivational role standardly ascribed to belief, but Velleman argues that it applies equally to the other attitudes involving acceptance, such as imagining and assuming. Consequently, belief cannot be characterized dispositionally, in terms of its 'output' side.[2] As far as the role in action explanations goes belief is not distinct from the other cognitive attitudes.

Since belief cannot be distinguished from the other cognitive attitudes simply at the level of the first order attitudes, Velleman argues, a 'two-tier structure' is required, appealing to the *aim* with which the first-order attitude of acceptance is formed. On this view, the first-order attitude of acceptance is combined with different second-order attitudes, "the different aims or intentions with which a proposition can be accepted" [Ibid., p. 113]. Whereas assuming involves accepting a proposition for the sake of argument, and imagining involves accepting it regardless of whether it is true, belief involves accepting *p* with the aim of thereby accepting the truth: the proposition "being accepted with that aim constitutes its being believed" [Ibid., p. 252].

This raises the question of what it means for an *attitude*, a psychological state, to have an aim. As first introduced by Velleman, the aim seems to belong to the agent: "...to believe a proposition is to accept it with the aim of thereby accepting a truth" [Ibid., p. 251]. Assuming, by contrast, involves accepting a proposition for the sake of argument. But this would be rather implausible. It seems clear that something can be a belief irrespectively of the individual's aims and intentions. Moreover, having a second-order attitude requires having the concept of the first-order attitude, and there are arguably individuals who have beliefs but do not possess the concept of belief (small children for instance). Velleman grants as much and says that his conception of belief does not require the aim of belief to be an aim on the part of the be-

liever. The aim, he suggests, can also belong to the cognitive system itself in the sense that it regulates cognitions "by forming, revising, and extinguishing them in response to evidence and argument" [Ibid., p. 253]. However, what matters is not merely that the system does in fact regulate cognitions in this way (that would allow for a characterization in terms of the first-order role) but that it was *designed* to do this: "If a cognitive state isn't regulated by mechanisms designed to track the truth, then it isn't a belief: it's some other kind of cognition. That's why aiming at the truth is constitutive of belief" [Ibid., p. 17].[3]

Velleman therefore takes the metaphor of belief aiming at truth as implying a teleofunctional account of belief according to which the essential feature of belief is the function of the mechanism that regulates it.[4] The relevant notion of a function is not one that can be understood purely causally (in terms of the causal power of the system), but is construed *teleologically*, in terms of what the cognitive mechanism is supposed to do.[5] This means that the relevant notion of function involves an appeal to the *history* of the system, what it was selected for. The two-tier account boils down to the claim that beliefs are a subclass of the attitudes of accepting *p* as true, i.e. those regulated by a mechanism designed or selected for truth:

> (T) Belief is an acceptance attitude regulated by cognitive mechanisms designed for tracking the truth.

Why should we accept (T)? The central assumption behind Velleman's account is the claim that the cognitive attitudes cannot be distinguished in terms of their first-order role. However, prima facie, it would seem that belief behaves very differently than imaginings and 'assumings', in practical reasoning as well as in theoretical reasoning. If so, there would be no need to adopt a two-tier structure – belief could simply be characterized in terms of its first-order role.

To illustrate this, consider Williams' further characterization of belief. Two central features of belief, according to Williams, are the fact that beliefs can be based on evidence and the fact that belief has an important connection with action. Both features concern the explanatory aspect of belief. To believe that *p* on the basis of evidence involves believing that *p because* of a further belief, the belief that *q*, which stands in an evidential connection to *p*: "if a man says to me 'Why do you believe that *p*?' I can rightly say 'because *q*'" [Williams (1973), p. 141]. The connection between the beliefs, Williams argues, is then both rational and causal. Thus, beliefs have a special role in theoretical reasoning: Beliefs are typically based on reasons and they cause further beliefs that they stand in rational connections with. Similarly, Williams argues, belief plays a special role in the explanation of action, in conjunction with the subject's projects or desires. If a man walks out on a bridge with de-

termined steps we can say that this shows that he believes that the bridge is safe, on the assumption that he desires to avoid drowning. This role that belief plays in action explanation in conjunction with desire, according to Williams, also serves to mark it off from other psychological states.

Velleman grants that it is typical of beliefs that they are responsive to evidence and argument. However, he argues, there are beliefs that are not responsive to evidence, for instance biased beliefs, and such beliefs fail to track truth, much like a phantasy. This is the extension problem, mentioned in the introduction, and Velleman responds to it by appealing to the aim of belief, its teleofunction. The difference between a belief and a phantasy, Velleman argues, does not lie in the behavior of the states but in the design of the mechanism that regulates belief: "Even when a belief is prevented from responding to corrective influences, the fact remains that its regulative mechanisms are being prevented from doing what they were designed to do. A phantasy and a biased belief are alike in that they fail to track truth; but the phantasy has no tendency to track the truth at all, whereas a biased belief is diverted from truth; and something can be diverted from truth only against the background of a tendency to track it" [Velleman (2000), pp. 254-255].

As noted above, Velleman also denies that belief plays a distinctive role in practical reasoning. To make his case, Velleman considers the phenomenon of *pretense* – children imagining being someone else, for instance the child imagining being an elephant [Ibid., pp. 256-258]. In such a situation, Velleman argues, the child is disposed to behave as if he were an elephant. It is not, as the standard belief-desire model would have it, that the child wishes to behave like an elephant and has beliefs about how to do that and acts accordingly. That would fail to capture the child-like quality of the pretense, make it too adult and calculated. Rather, the child enters the fiction and acts it out, and this involves acting *out of* the imagining. The motivating cognition is not 'Here is how to behave as if this chair were a pail of water' but, simply, 'Here is a pail of water'. The latter thought is not a belief, Velleman argues, but an *imagined* belief, which operates, and motivates, as if it were a real belief [Ibid., p. 259]. Similarly, the child will have 'mock-desires', for instance the 'desire' to take a drink, that play the motivational role of desire. The person who pretends therefore acts out of motives she imagines herself to have:

> In the fiction that I am an elephant, my imagining and wishing are a belief and desire, moving me to drink from the pail. When my imagining and wishing move me to behave as if drinking, they fulfill the motivational role of the belief and desire that they are imagined to be, with the result that I enact my imagined role as an elephant [Ibid, p. 261].

Velleman is not alone in thinking that imagining that *p* plays a motivational role very much like belief. For instance, Tyler Doggett and Andy Egan have

argued that imagining that *p* (but not supposing that *p*) motivates us to act directly, in conjunction with 'i-desires', an analogue to ordinary desires [Doggett and Egan (2007)].[6] When pretending to be a cat, they argue, the subject need not have any specific beliefs about how cats behave; rather, she imagines being a cat and then this imagining moves her to act a certain way. Like Velleman they hold that this is required in order to account for imaginative immersion, the capacity children (and some adults) have of 'losing themselves' in the pretense. Thus, they claim, "pretense-directed behavior is generated by imagination in very much the same way as regular, non-pretense directed behavior (...) is generated by belief" [Ibid., p. 9].[7]

Now, in order for these considerations to count against the thesis that belief can be characterized in terms of its motivational role, the claim cannot merely be that imagining plays *a* role in the explanation of pretense behavior. Rather, the claim must be that the attitude of imagining plays the *same* motivational role as belief. And this would seem to be a rather problematic claim. Belief is regularly connected with action, whereas imagining is not: Most of the time, we do not act on our imaginations, and when we do this would seem to be motivated by the specific desire to 'act out' our imaginations.[8] Velleman recognizes that imagining is less likely to cause actual behavior than beliefs are, but he explains this as a combined result of 'countervailing beliefs' and inhibitions. First, he argues, imaginings are accompanied by countervailing beliefs "embodying the subject's knowledge of the facts that he is imagining to be otherwise, such as his knowledge that an imagined pail of water is really a chair" [Velleman (2000), p. 272]. Beliefs are not normally accompanied by such countervailing beliefs, Velleman suggests, and for this reason their motivational force is not in the same way blocked. Second, the motivational force of imaginings comes under an inhibition acquired on the way to adulthood. We learn to separate our realistic conations and cognitions from the unrealistic ones, and to suppress the motivational force of the unrealistic ones [Ibid., pp. 262-263].

However, it does not seem to be an accidental feature that imaginings, but not beliefs, are accompanied by 'countervailing beliefs'. As Velleman himself suggests, when the child thinks 'This is a pail of water' of what she knows is a chair, this thought cannot be a belief since one cannot believe that something is a pail of water if one knows that it is a chair. That is, one cannot (in normal circumstances) believe that *p* when one knows that not *p*.[9] By contrast, the child can *imagine* that the chair is a pail of water even if she knows that it is a chair; indeed, it would seem that imagining that the chair is a pail of water requires believing that it is not a pail of water. This reflects a clear sense in which belief is 'truth-directed' whereas imagining is not: While there is no obstacle to imagining that *p* when one knows that *p* is false, only in the exceptional case can one believe that *p* when one knows that *p* is false. This, also, explains why imagining is 'less likely to cause actual behavior':

To imagine that *p* does not (normally) 'dispose one to behave as if *p* were true', since imagining that *p* (normally) involves believing that not-*p*. To be disposed to behave as if *p* were true, when one knows that *p* is not true, one has to have the desire to act out one's imaginations, to act *as if p* were true.[10] This difference between belief and imagining is related to the phenomenon of *quarantining* that has been suggested to be distinctive of pretense.[11] When one pretends that a chair is a pail of water, things that are believed to be true of water do not thereby come to be believed true of the chair. For instance, imagining the chair to be a pail of water the child does not expect the chair to contain a wet and thirst-quenching liquid. Imagining that *p* is in this sense inferentially isolated, in a way that ordinary beliefs are not: If I believe of an object that it is a pail of water I will indeed expect it to contain a wet, thirst-quenching liquid. This, also, points to an important difference in the motivational role of belief and imagination. Whereas beliefs interact with further beliefs to produce actions, imaginings do not. If I imagine that I am an eagle, and I desire to fly, I am not thereby motivated to jump off a cliff.

When it comes to Velleman's second point, about the role of inhibitions, it is no doubt true that adults are generally speaking worse at pretending than children precisely because of acquired inhibitions. However, it does not seem true that it is a matter of learning to separate realistic attitudes from unrealistic ones. Developmental evidence shows that when children first learn to pretend (at around fifteen months), they have clear awareness of the pretend status of what they are doing. A fifteen-month old child who pretends that pieces of cloth are pillows shows clear awareness of the pretend status of the cloth.[12] The phenomenon of quarantining therefore seems to be there from the start, as an essential feature of the pretense.

We shall return to the role of belief in action explanation below. These preliminary remarks do suggest, however, that Velleman is wrong to claim that imagining that *p* plays the same motivational role as believing that *p*. If so, it remains a live possibility to appeal to the role of belief in the production of action as one of its distinctive marks, and to develop an account of belief without appealing to the design of the belief regulating mechanism. Next, we shall suggest that this would be a welcome consequence, since the teleofunctional account implies a problematic separation of the nature of belief from its first-order role in practical and theoretical reasoning.

III. PROBLEMS FOR TELEOFUNCTIONALISM – AND NORMATIVISM

The claim that the design of the regulating system is essential to belief has two problematic, and related, implications: First, it follows that it is a necessary truth that a creature whose cognitive system was not designed for

truth does not have beliefs. Second, it follows that if a particular state is not regulated by a mechanism designed for truth, it is not a belief.

The first problem is closely related to the difficulties facing Davidson's account of the 'swampman': An individual may behave exactly like us, like a rational agent and full-fledged language user, and yet not have beliefs simply because of its history [Davidson (1987)].[13] Indeed, we know very little about the design of our own cognitive systems – what if the system that regulates what we take to be beliefs were not designed for truth? According to Velleman's account, it would follow that we do not have any beliefs.

Velleman is aware of the difficulty, but claims that it is extremely improbable that belief could be designed for something other than truth. This, he suggests, can be discovered through introspection: "When we discern a gap between a belief and the truth, the belief immediately becomes unsettled and begins to change. If it persists, we form another belief to close the gap, while reclassifying the recalcitrant cognition as an illusion or a bias. I cannot imagine evidence that would show this reclassification to be a mistake" [Velleman (2000), p. 278]. This seems largely accurate. It is obscure, however, how this observation could tell us anything about the *design* of the mechanism regulating belief. As far as the design is concerned, it is entirely conceivable that belief behaves the way it does, while the system in fact was designed for something else, such as instrumental success.[14]

The second problem concerns the existence of irrational beliefs, or the extension problem. As noted above, one reason Velleman adopts a teleofunctional account is precisely to solve this problem. However, it would seem that in the case of many irrational beliefs, such as wishful thinking and biased belief, the belief in question is not regulated by mechanisms designed for truth. If so, Velleman would have to say that such attitudes could not be beliefs. Velleman notes that this raises a difficulty for him. Thus, he mentions the case of delusional beliefs, as when someone believes he is Napoleon, which would seem to suggest that we could have beliefs that are not regulated for truth. Velleman suggests, however, that a delusional belief cannot literally be said to be a belief. When we say that someone believes he is being Napoleon we are only using the term 'belief' in a figurative sense: "The phantasy of being Napoleon is thus what he has instead of a belief about his identity" [Ibid., p., 281].

It is a good question how irrational a state can be and yet be classified as a belief. Even if a plausible account of belief must allow for the possibility of irrationality, intuitively there must be some limits to the irrationality allowed for the state to qualify as belief. Depending on the details of the case, it may be that delusional 'beliefs' of the sort Velleman discusses fail to qualify. However, it does not seem right to suggest that the difference between an irrational belief and a non-belief derives from the mechanism that produces it: There may well be cases of irrational beliefs, such as wishful thinking or

biased beliefs, which are produced by protective mechanisms having little to do with truth. If, therefore, Velleman is committed to saying that beliefs involved in wishful thinking and bias are not beliefs, teleofunctionalism remains too restrictive to (fully) solve the extension problem.

One common response to this difficulty is to suggest that the link between belief and truth is *normative*. On this view what is constitutive of belief is not that belief is designed to behave a certain way, or that it does behave a certain way, but that it *ought* to behave a certain way: Belief is that state which is governed by the truth norm.[15] The truth norm comes in various versions, the strongest one being:

(N) One ought to believe that *p* if and only if *p*.[16]

Since the relation between truth and belief is merely normative, it is possible to have a belief that is not responsive to truth, such as in cases of bias or wishful thinking. Indeed, on a purely normativist account of belief a state could be a belief quite independently of how it in fact behaves. Although my belief that *p* ought to be revised when I come to believe that *not-p* this may or may not happen, and although I ought not to believe that which is unsupported by evidence it may be that all of my beliefs lack evidential support: As long as the state in question is governed by the relevant norm, it is nevertheless a belief.

We have argued against normativist accounts of belief elsewhere [(2009, 2010, 2013)]. Here we would like to stress an additional problem with normativism, one that it shares with teleofunctionalism: Both the teleological account and the normativist account imply a problematic disconnect between the nature of belief and its first-order role. In both cases, the disconnect is a reaction to the extension problem, the fact that although beliefs typically are evidence-sensitive, some beliefs are formed on insufficient evidence and, in extreme cases such as wishful thinking, in the face of strong counterevidence. The teleofunctional response, again, consists in appealing to the design of the mechanism that regulates the state, in conjunction with the idea that such a mechanism may fail its purpose: If the state is regulated by the right mechanism it is a belief, even if it is not evidence responsive. The normativist response consists in making the connection with truth prescriptive: Beliefs should be formed on the basis of evidence, etc. but they may not.

This means that both teleofunctionalism and normativism face the swampman problem: The former, since it follows that swampman may have cognitive states that behave like beliefs (by and large responsive to evidence, etc.) but which are not beliefs since they are not regulated by a system designed for truth; the latter since it follows that swampman may have cognitive states that behave like beliefs but which are not beliefs since the subject is not guided by the truth-norm (or any related norms, such as the norm of

sufficient evidence). Conversely, both theories imply what might be called the "wild man problem": It follows that a wildly irrational subject may have beliefs, states that do not stand in any evidential or logical connections to one another (and do not serve to rationalize action) and yet are beliefs. Teleofunctionalism (arguably) has this consequence, since one can imagine the subject's cognitive mechanism malfunctioning radically; normativism has this consequence since from the fact that the subject ought to form beliefs in a certain way, nothing follows immediately about his actual belief formation.[17]

We take this divorce between the first-order role of belief and its essence to be unacceptable. The intuition underlying William's metaphor of belief aiming at truth, we submit, is not that belief ought to aim at truth or that it is designed to aim at truth, but that belief in fact (by and large) aims at truth. Indeed, as noted above, this is recognized on all sides of the debate. Both those appealing to teleological function and those appealing to norms tend to add that there is a *descriptive* element involved such that belief is that state which, by and large, is responsive to evidence and stands in inferential relations to other beliefs. For instance, in their joint paper from 2005, Velleman and Shah defend a hybrid account of belief, according to which the concept of belief "is that of a cognition that is governed, both normatively and descriptively, by the standard of truth" [Velleman and Shah (2005), p. 499].[18]

However, falling back on a descriptive component merely shows the normative element to be redundant. If indeed belief has a distinctive first-order role, there is no need to appeal to norm guidance in order to separate belief from the other normative attitudes. Similarly, there is no need to appeal to the etiology of the belief regulating mechanism, it is sufficient to appeal to its actual (non-teleological) functional role. At points, in fact, Velleman suggests as much in his paper on the aim of belief. Thus, he argues that the 'reality-tested' quality of belief serves to mark it off from the other attitudes involving regarding *p* as true and, similarly, serves to mark off desires from the other conations:

> Among the thoughts that we are disposed to make true – that is, among our cona-
> tions – we delimit a subset whose members we are disposed to revise, discard, or
> at least reclassify if we cannot actually make them true. These reality-tested co-
> nations are our desires, which interact with one another in relative isolation
> from our mere hopes and wishes. Similarly, among the thoughts for which we
> have a disposition to behave as would be desirable if they were true – that is,
> among our cognitions – we delimit a subset whose members we are disposed to
> revise, discard, or at least reclassify if they aren't actually true. These reality-
> tested cognitions are our beliefs, which interact with one another in relative iso-
> lation from our mere imaginings [Velleman (2000), p. 263].

To say that beliefs (and desires) are 'reality-tested' is precisely to accept the idea that beliefs and imaginings play very different roles (and that desires and

mere wishes play very different roles). It is to suggest that belief can be characterized without appealing to its teleofunctional aim, in terms of its first-order role in theoretical and practical reasoning. [19]

The lesson of this, we take it, is that we should not attempt to solve the extension problem by divorcing the essence of belief from its first-order role, as both normativism and teleofunctionalism do. Next, we shall consider another reaction to the extension problem, one that ties the nature of belief very closely to first-order role. On this view, it is simply constitutive of belief that it is rational, and any state that behaves differently must be classified as some other type of cognitive attitude.

IV. ALIEF AND ACTION-EXPLANATION

In a series of important recent papers, Tamar Gendler has introduced and developed the notion of a hitherto unrecognized cognitive mental state: That of *alief* [Gendler (2008a, 2008b)]. Without a notion like alief, she argues, we are bound to seriously mischaracterize not only what it is to have beliefs, but also and more generally what it is to be motivated to act: "The problem with the belief-behavior picture is that at its heart lies a faulty picture of what makes us act" [Gendler (2008a), p. 655]. Intriguingly, according to Gendler we need alief, and the new picture of action motivation that comes with it, precisely in order to save the traditional philosophical conception of belief:

> I will argue that any theory that helps itself to notions like belief, desire, and pretense needs to include a notion like alief in order to make proper sense of a wide range of otherwise perplexing phenomena. (...) In short, I will argue that if you want to take seriously how human minds really work, and you want to *save belief*, then you need to make conceptual room for the notion of alief [Gendler (2008a), pp. 641f, emphasis added].

Gendler's papers contain rich descriptions and detailed discussions of the "otherwise perplexing phenomena" she is rightly interested in making sense of. We cannot do justice to any of that here. We are after belief, and our present interest in alief is limited to its potential for saving belief. More precisely, our question is whether the notion of alief can be used to provide a first-order solution to the extension problem. We shall therefore be quick in introducing the notion of alief, and restrict discussion to those of its features that are relevant for our question.

According to Gendler, alief is omnipresent in our cognitive lives. Its presence is most easily recognized, however, when it 'acts up'. She therefore introduces her readers to the notion by means of a certain kind of example.

These are examples of what she calls "belief-discordant alief".[20] Here are two: In what we shall call "the cliff case", a subject tries to walk onto the Grand Canyon Skywalk, a semi-circular glass bridge hanging over the edge of the canyon. He is completely convinced that the bridge is perfectly safe. Nevertheless, he is trembling and anxiously recoils. The reaction is so strong that the subject does not manage to walk out on the bridge. The second example, let's call it "the poison case", involves experiments with subjects who see two glass bottles being filled with sugar from the very same box. Then, they themselves label one of the bottles 'sugar' and the other 'sodium cyanide'. And subsequently, they show reluctance to consume the sugar from the second bottle.[21]

In cases such as these, the subjects are quite naturally described as acting contrary to their beliefs (and desires), or as at least finding it perplexingly difficult to act in accordance with them. Moreover, the tendencies or dispositions manifested can be extremely hard to eradicate. In particular, they can be almost completely immune to countervailing evidence. According to Gendler, trying to construe the cognitive states underlying and explaining the subjects' behavior in these cases as beliefs is therefore deeply misleading. What we need is a second main category of cognitive mental state, a mental state like belief in being action motivating, but unlike belief in being largely evidence-immune: Alief.

"An alief is, to a reasonable approximation, an innate or habitual propensity to respond to an apparent stimulus in a certain way", Gendler explains [(2008b), p. 553]. She offers the following, more precise "tentative characterization of a paradigmatic alief":

> A paradigmatic alief is a mental state with associatively linked content that is representational, affective and behavioral, and that is activated – consciously or nonconsciously – by features of the subject's internal or ambient environment. Aliefs may be either occurrent or dispositional [Ibid., p. 642].

Aliefs are mental states with *content*, but they are *not* propositional attitudes. This is for two reasons: First, Gendler uses 'content' in a "somewhat idiosyncratic way" [Ibid., p. 635, fn. 4]. The content of an alief consists of three associatively linked components: A representation R "of some object or concept or situation or circumstance" [Ibid., p. 643], an affective component A, and a behavioral component B described as "the readying of some motor routine" [Ibid.]. Therefore, alief is not a two-place relation between a subject and a proposition, but a four-place relation between a subject and components R-A-B. Second, the R component does not have to be a proposition, according to Gendler. The content of the alief in the canyon case could thus roughly be given like this: "Really high up, long long way down. Not a safe place to be! Get off!!" [Ibid., p. 635].

Aliefs have representational content, but they are *not* attitudes of acceptance: "Unlike belief or pretense or imagination or supposition, alief does not involve acceptance" [Ibid., p. 648]. Moreover, Gendler comments, the notion of representation used in characterizing alief "is a thin one: because they involve mechanisms that are wholly insensitive to the difference between seeming and being, or between appearance and reality, aliefs lack certain sorts of correctness conditions" [Ibid., 559, fn. 11]. This, however, is more a comment on the kind of representational *state* alief is than on its content, it seems to us. Moreover, it is entirely in line with the in our context most interesting contrast, the contrast between aliefs and beliefs:

> Aliefs by their nature are insensitive to the possibility that appearances may misrepresent reality, and are unable to keep pace with variation in the world or with norm-world discrepancies. By contrast, beliefs are (modulo error) responsive to the way things are: not merely to the way things tend to be or to the way things seem to be [Ibid., p. 570].

In other places, Gendler makes the contrast in terms of reality-sensitivity: "Alief is just not reality-sensitive in the way belief is. Its content does not track (one's considered impression of) the world" [Ibid., p. 651]. Or in terms of evidence-sensitivity: "Beliefs change in response to changes in evidence; aliefs change in response to changes in habit. If new evidence won't cause you to change your behavior in response to an apparent stimulus, then your reaction is due to alief rather than belief" [Ibid., p. 566].[22]

It should be fairly clear, however, that the contrast needs to be made more precise. As the alief-belief distinction is formulated here, certain states that intuitively pretty clearly are beliefs would fall on the alief side. Huddleston (2012) has some relevant examples of what he calls "naughty beliefs": states that are intuitively beliefs, but quite as evidence-immune as the phenomena Gendler describes – such as certain paranoid or phobic beliefs. Gendler is aware of this; in a footnote, she writes: "As stated, this principle is too strong, for there are certainly cases of subjects who hold evidence-recalcitrant *beliefs*" [Ibid., p. 566, fn. 26]. In these cases, she suggests, subjects somehow distort the evidence, but admits that the response is preliminary at best. At this point in the discussion, it therefore remains open whether the alief-belief distinction will be able to solve the extension problem. As just illustrated, there is a worry that – just like Velleman's teleofunctionalism – it will remain too restrictive and exclude states from the realm of belief that intuitively belong there.

In the remainder of this section, we would like to dwell on another worry, however. This worry concerns the similiarities between alief and belief more than the differences: The question is whether alief and belief can be as

similar as Gendler needs them to be without the distinction ultimately collapsing.

Alief is *similar* to belief in its output role: both motivate and explain behavior. Crucially, alief-motivated and alief-explained behaviors are not 'just behavior', not just events or movements happening to their subjects' bodies. Rather, they are either *intentional actions* or at least *sufficiently similar to intentional actions* to allow for explanation by means of cognitive states with representational content.[23] Given how different from beliefs aliefs are supposed to be, this would mean that the traditional picture of action and action explanation we sketched in the introduction has to be significantly modified.

It would have to be enriched by an alternative way of explanation:

> Introducing the notion of alief into our descriptive repertoire provides a useful alternative way of answering 'why?' questions when confronted with a behavior or tendency that we seek to explain. It provides *an alternative that falls somewhere in between a classic reason-based explanation* (of the sort offered by belief/desire accounts) *and a simple physical-cause explanation* [Gendler (2008b), p. 555, emphasis added].

Alief-explanation is supposed to be sufficiently similar to reasons explanation to be recognizable as a kind of *intentional explanation in a wide sense*. By 'intentional explanation in a wide sense' we mean an explanation in which the content of the mental states cited in the explanation is not explanatorily idle. Rather, the explanatory force of such an explanation depends both on the states' having content, and on *which* content they have. At the same time, however, alief-explanation needs to remain sufficiently different from intentional explanation in the narrow sense of belief(-desire)-explanation – otherwise the notion of alief would just collapse into that of belief. The crucial question here is thus whether there indeed is room for another notion of intentional explanation (in a wide sense) between belief-desire explanation and purely causal explanation. Our worry is that there isn't. There seems to be a dilemma here, a dilemma we shall call "the alief-dilemma": Either alief-explanation is recognizably intentional. Then alief becomes too much like (irrational) belief. Or alief is sufficiently different from (irrational) belief. Then alief-explanation is no longer recognizably intentional – not even in a wide sense.

Before we set out the dilemma, a couple of words about the picture of mental content we are employing. Contents are abstract objects assigned to certain kinds of mental states to model their role in our psychology. In order to provide a good model, these abstract objects need to be of a certain kind. They need to be sufficiently fine-grained to capture all the distinctions we want to capture, and they need to be related to one another in ways apt to capture all the relations between the states we are interested in capturing.

That's why propositions (whatever they precisely are) are better for modeling propositional attitudes than, say, numbers.[24] In what follows, we shall work with this picture of content. We shall also work with a 'minimal' notion of proposition: Something is a proposition iff it essentially has conditions of truth, correctness, or accuracy.

On this picture, assigning propositional contents to certain mental states allows for modeling theoretical and practical reasoning by means of (logical or 'material') inferential relations between propositions. One such state can provide a reason for another, or for acting a certain way, precisely in the sense of there being a valid inference from one content, or set of contents, to another. This is the sense in which intentional explanation requires the mental states it cites to have content, and it is also why it matters to the explanation which content those states have. Our question thus is: Is alief-explanation sufficiently similar to this picture of intentional explanation to be recognizably intentional in a wide sense?

According to Gendler, alief does have representational content, but this content does not need to be propositional. We are not completely sure what she takes to be required for a content to be propositional, but assuming a minimal notion of propositional content, it is hard to see how alief could lack such content. Its *R*-component is supposed to represent such things as a steep cliff, or a bottle of poison. Moreover, it is plausible to assume that it represents that cliff as in front of my feet and a bottle of poison as present. Otherwise, it would be hard to see how having such 'content' could explain anything at all. Just 'representing' a cliff that could be anywhere, or a bottle of poison that might not even exist, – would not plausibly move anybody.[25] Think of such a 'content' on the model of an open formula '... is a bottle of poison' or a free-standing singular term 'o ...'. Ascribing such partial contents does not help explain behavior in a sense where the content matters. Having such partial contents is not like seeing a sign reading 'Dangerous cliff' placed right on its edge. Rather, it is like finding part of a torn piece of paper with the word 'cliff' on it in a drawer. Just by itself, representing this concept or property does nothing to motivate or explain acting in any particular way. Of course, tokening representations with partial contents could have all sorts of causal effects, but that was not the kind of explanation we were after. So, it seems to us, the only kind of content that can make the right kind of difference to the explanation of behavior is propositional content.

Things get worse, however. Merely representing a propositional content does not explain anything, either. You can entertain all sorts of propositions, but such entertaining, just by itself, will not move your body. A lively imagination might get you closer to action, but – just by itself – imagining a proposition to be true won't move your body, either.[26] Imagination points in the direction of what is needed, however: A certain relation to truth. In general, we can distinguish between two senses of representing or representational

state: A weak sense, in which a state is representational iff it has a proposi-
tional content *p*. And a strong sense, in which a state is representational iff it
has a propositional content *p* and represents the world as being such that *p*
obtains.[27] To motivate action, it seems to us, it is not enough for a state to be
weakly representational. It needs to be strongly representational.

The point can be made in terms of content, again. Even if a state has a
truth-conditional content, this content needs to be connected to the *actual*
world to explain or motivate behavior. Just by itself, representing a way the
world might, or might not, be does not motivate or explain acting in any par-
ticular way. Of course, merely tokening representations with truth conditional
contents could have all sorts of causal effects, but again, that was not the kind
of explanation we were after. But what connects a truth-conditional content
to the actual world, what makes a state into a strongly representational state is
not the content itself, but the state's 'mode' or attitude component. Being
strongly representational thus is being a certain kind of attitude, even a cer-
tain kind of propositional attitude. Consequently, it seems to us, the only kind
of attitude that can make the right kind of difference to the explanation of be-
havior is a strongly representational or 'committal' attitude.

Gendler seems to deny that, too. At least, she denies that alief involves
acceptance. But the phenomena she cites in her argument against alief being
an acceptance attitude actually do not support this conclusion. As far as we
can see, Gendler is perfectly free to construe alief as an acceptance attitude.
She bases her denial of this on the results of another of Rozin's poison bottle
experiments. In this experiment, the labels the subjects apply to the bottles
containing what they perfectly well know to be sugar read 'sucrose, table
sugar' and 'not sodium cyanide, not poison', respectively. Even in this ex-
periment, subjects were more hesitant to consume sugar from the second bot-
tle, though the effect was weaker than in the original experiment [cf. Gendler
(2008a), pp. 649f]. Gendler's explanation of this would of course be that
though the subjects believed that the bottle did not contain sodium cyanide or
any other poison, they alieved that it did. But why would this show that the
alief was not an acceptance relation to its own content? Gendler explains:

> [A]lthough these subjects were in an alief state with the content "cyanide, dan-
> gerous, avoid", the content they were prompted to imagine was exactly the op-
> posite. They did not – as the acceptance condition requires – regard it as true in
> some way that cyanide is to be found in the vicinity; instead, it was the negated
> presence of the word "cyanide" that rendered occurrent their cyanide associated
> aliefs [Ibid., p. 649].

And a little later, she writes: "At its core, alief involves the activation of an as-
sociative chain—and this is something that can happen regardless of the atti-
tude one bears to the content activating the associations" [Ibid., p. 650]. But
that seems to rest on identifying a content that is the stimulus activating an alief

with the content of the alief itself. There is no reason for this identification, it seems to us. Of course, the label prompts the subject to entertain or imagine the content that the bottle does not contain cyanide. And the presence of the word 'cyanide', even in negated form, clearly has powerful psychological effects that we can describe as activating an alief. But that in no way forces the conclusion that if alief were an acceptance relation, the accepted content would be that the bottle does not contain cyanide. Even if the subject entertains, imagines, and of course believes that content, it is perfectly coherent and – we think – plausible to think of the alief as acceptance of the content that the bottle contains cyanide.[28]

In sum, it seems to us that aliefs not only need to be attitudes to propositions for alief-explanations to be recognizably similar to core cases of intentional explanation. They also need to be strongly representational or acceptance attitudes. Otherwise, it is just not clear how their having 'content' makes for the right kind of difference between an alief-explanation and a merely causal explanation. But if alief is an acceptance relation to a proposition, is alief still sufficiently different from belief?

To be sure, alief still requires association with an *A* and a *B* component.[29] And it would need to be highly 'modular' both in order to retain the degree of evidence-immunity characteristic of it and in order to prevent it from exerting undue influence on those of its subject's states that clearly are beliefs. Nevertheless, alief has become *very* similar to belief, so similar that we are not sure anymore that alief really is a mental state sui generis. Belief-discordant aliefs now rather look like beliefs of a highly compartmentalized subject. Many of these are plausibly described as irrational, whether categorized as alief or as belief, and many of these are plausibly associated with strong emotional reactions, again regardless of whether they are categorized as aliefs or as beliefs. Paranoid or phobic beliefs are good examples here.

Now, we are *not* claiming that all belief-discordant alief in fact is better construed as compartmentalized belief. There might well be phenomena here where such a construal remains utterly implausible. However, the more implausible such a construal becomes, the less plausible it is to construe the resulting behavior as intentional action – in any sense. That is the alief-dilemma: Either alief-explanation is recognizably intentional. Then alief becomes too much like (irrational) belief. Or alief is sufficiently different from (irrational) belief. Then alief-explanation is no longer recognizably intentional – not even in a wide sense. If this is correct, then the traditional notion of belief needs to be rescued some other way.

V. CONCLUSION

We have argued that teleofunctionalism and normativism about belief allow implausibly much divergence between belief's first-order functional

role and the norm or *telos* they construe as essential to belief. This provides motivation for renewing the attempt at accounting for belief precisely in terms of its first-order functional role. To the extent that that involved characterizing belief as an evidence-sensitive propositional attitude, any such account has to deal with what we have called the extension problem: Actual belief formation often falls short of being optimally evidence-sensitive, and some belief might even be (almost) completely evidence-immune. Trying to solve the extension problem by means of introducing the notion of alief is intriguing, but as worked out so far, it remains unclear whether alief can be made to do the work required for 'saving belief' – without collapsing into belief.

The main ways of construing belief's aiming at the truth currently on the market thus seem to lead into dead ends. We would like to take an optimistic stance on this diagnosis. We would like to think that this clears the way for an even more traditional functionalism about belief, rather than supports the conclusion that functionalism itself is at a dead end. Whence the optimism? Well, none of the positions we have looked at makes use of the *whole* functional role of belief. On the input side, recent discussion stresses evidence-sensitivity, and on the output side, motivation. What is neglected is the role of belief in theoretical, as opposed to practical reasoning, especially its output side. But – or so we claim – belief has a characteristically truth-directed reason-providing role not only for action, but also for further belief.[30] Therefore, the hope for characterizing belief in terms of its distinctive evidence sensitive, reason providing role for both action and further belief is not out yet.

Department of Philosophy
Stockholm University
10691 Stockholm, Sweden
E-mails: kathrin.gluer@philosophy.su.se
*   asa.wikforss@philosophy.su.se*

NOTES

[1] We have made a first stab at developing our version of Davidsonico-Lewisian functionalism in a presentation given at the "Aim of Belief" workshop in Oslo, June 2009. A podcast of the presentation is available at *http://www.csmn.uio.no/podcast/pagin_wikforss.html*.
[2] For a defense of the idea that belief can be characterised dispositionally see for instance Williamson (2000). Williamson argues that the difference between believing *p* and merely fancying *p* "depends in part on one's dispositions to practical reasoning and action manifested only in counterfactual circumstances" [p. 24].

[3] It is a good question whether the teleofunctional account of the 'aim' of belief could be extended to the other cognitive attitudes. Velleman's claim, again, is that each such attitude is distinguised by the aim with which the proposition is accepted as true. Presumably, there will be no evolutionary story forthcoming in the case of imaginings and assumptions. If so, Velleman must hold that in these cases the aim belongs to the subject: Imagining that *p* is accepting *p* as true with the intention of doing so regardless of its truth. This means that one cannot imagine that *p* without having quite sophisticated intentions.

[4] For a similar idea see Papineau (2013). Papineau suggests that from the point of view of biological design the function of belief is to help select actions that will produce satisfaction of desire. To have this function beliefs need to carry information about the environment and in this sense beliefs 'aim at truth'.

[5] For the distinction between the etiological notion of a function and the purely causal, non-teleogical notion (so-called Cummins functions), see Häggqvist (2013).

[6] See also Currie and Ravenscroft (2002).

[7] Unlike Velleman, however, Doggett and Egan do not take i-desires to be a type of wanting, but suggest that i-desires constitute a separate psychological category: Imagining that you are a cat you may i-desire to keep your paw clean and this may move you to lick your hand, but this does not mean that you *wish* to keep your paw clean [Doggett and Egan (2007), p. 11]. This is connected with the fact that Doggett and Egan hold that there is a difference in the functional role of i-desires and that of ordinary desire, and they also suggest that there is a difference in the functional role of imagination and that of belief [Ibid., p. 10]. This may be one reason they do not appeal to Velleman's teleofunctional account of belief.

[8] For a discussion see Funkhouser and Spaulding (2009), Gendler (2003), O'Brien (2005) and Zalabardo (2010). They all argue that while imagining may play some motivational role, this role is very different from that of belief. For a belief-desire account of pretense see also Nichols and Stich (2003).

[9] There may be extreme psychological circumstances (such as cases of serious paranoia) under which a subject believes *p* (in full awareness) while she knows that not *p*. See Huddleston (2012) for discussion. We return to this briefly below, in section IV.

[10] See O'Brien (2005) who stresses that such a desire is required. She also argues that acting out one's imagination requires having certain beliefs about how one can use one's body and environment to act out the imagination (such as the belief that one can use the arm as a pretend-trunk, and the chair as a pretend pail of water). This, she suggests, shows that imagining that *p* does not play the same motivational role as the belief that *p*, but that imagining piggy-backs on the world-directed attitude of belief. [O'Brien (2005), p. 58].

[11] See Gendler (2003), pp. 130-131. Another distinctive feature of pretense discussed by Gendler is the principle of *mirroring*, according to which what is believed affects what is pretended. For instance, if I pretend that the chair is a pail of water I will also pretend that it contains a thirst-quenching liquid. This illustrates the point stressed by O'Brien, that we exploit our beliefs about the world in acting out our imaginations. And the principle of quarantining tells us that the converse does not hold: We do not form beliefs on the basis of what is pretended.

[12] See Gendler (2003), for a discussion of the relevant literature.

[13] See also Dretske (2000) and Millikan (1993) for the claim that the history of the mind is essential to it. Davidson, Dretske and Millikan all endorse this claim on the basis of considerations having to do with the determination of mental *content*, rather than with the nature of belief. However, the swampman problem arises in both cases (although, of course, the theories of content have more disastrous consequences since it follows that swampman cannot have *any* intentional states). See Häggqvist (2013), for a discussion.

[14] This problem is stressed by Zalabardo (2010). He notes that even if the system that regulates belief mostly produces true beliefs "this fact might be, from the evolutionary point of view, nothing but a happy accident" [p. 18]. Zalabardo also suggests an alternative interpretation of the teleofunctional account, according to which 'belief' is a natural kind term such that we can discover its essential nature only a posteriori. This avoids the difficulty, since if we discover that belief is not designed for truth it would not follow that there are no beliefs but merely that the essential nature of belief is different from what we assumed it to be. However, as Zalabardo notes, the theory runs into a related problem since it implies that in a counterfactual world where our cognitive states were not designed for truth, we could not have beliefs – even if we had states that were functionally equivalent to beliefs.

[15] This is the line defended by Nishi Shah (2005), pp. 461-465. In a joint paper from the same year, Shah and Velleman endorse a hybrid account of belief, according to which belief is to be characterized both teleologically and normatively.

[16] For a discussion of the form of the truth norm see Bykvist and Hattiangadi (2007).

[17] Whether the teleofunctional account leads to the wild man problem depends on the details of the underlying account of teleological functions [for a discussion see Häggqvist (2013)]. But if one allows that a heart can malfunction utterly and still be a heart, the same would presumably hold for the belief regulating mechanism.

[18] See also O'Brien who suggests that an account of belief in terms of its motivational role needs to be supplemented by a normativist account according to which belief involves being responsive to norms of rationality [O'Brien (2005), p. 62].

[19] It is sometimes said that the causal role essential to belief includes relations to non-intentional states such as feelings and emotions. For instance, José Zalabardo argues that it is distinctive of belief to be disposed to cause the conscious state of feeling that a proposition entertained is true. Zalabardo suggests that this is the most obvious difference between beliefs and assumings, since when I assume that *p* there will be no feeling of conviction [Zalabardo (2010), p. 16]. It is entirely possible that, in many cases, there is a phenomenological difference of this sort between believing that *p* and assuming that *p*. It is doubtful, however, that this can be said to be essential to belief, since it implies that there cannot be a creature who has beliefs without such a phenomenology.

[20] She also calls the aliefs manifested in these cases "norm-discordant" because they violate "certain norms of cognitive-behavioral coherence" [Gendler (2008a), p. 651]. It is not completely easy to square that with the further claim that alief is *a*rational, however, [cf. Ibid., p. 641]. The norms violated by belief-discordant alief presumably are norms of rationality. States subject to norms of rationality can be irrational, but hardly arational.

²¹ Cf. Gendler (2008a), p. 648. The experiments are from a 1986 study by Paul Rozin.
²² These passages all sound as if the contrast Gendler is after were purely descriptive in nature. And that is how we shall take her here, mainly, but not only, because it is in this form that her proposal is most interesting for our discussion. There are, however, places where she seems to be describing the contrast between alief and belief in normative terms, as when she writes: "Whatever belief is – it is normatively governed by the following constraint: belief aims to 'track truth' in the sense that belief is subject to immediate revision in the face of change in our all-things-considered evidence"[Gendler (2008b), p. 565]. Note that even formulations like the one just quoted describe belief as *being* subject to immediate revision in the light of the strongest evidence, *not* as what ought to be so revised. At the core of Gendler's argument, it therefore seems to us, *is* a purely descriptive contrast: Aliefs are not just wayward beliefs, because aliefs in general behave so differently from paradigmatic beliefs. More precisely, they are formed in ways very different from the way typical ways of belief formation. For more discussion, see also Nagel (2012).
²³ Compare and contrast the subject recoiling from the Skywalk with the climber in Davidson's example of a deviant causal chain between a belief, a desire, and what is precisely *not* an intentional action, but something merely happening to the subject of the belief and desire:

> A climber might want to rid himself of the weight and danger of holding another man on a rope, and he might know that by loosening his hold on the rope he could rid himself of the weight and danger. This belief and want might so unnerve him as to cause him to loosen his hold, and yet it might be the case that he never chose to loosing his hold, nor did he do it intentionally [Davidson (1973c), p. 79].

We are not completely sure whether Gendler thinks of alief-motivated behavior as intentional action or merely as in between intentional action and mere event. In one place, she writes: "Belief-desire explanations are supposed to explain (or 'rationalize') *intentional* actions – not mere behaviors. But of course, that's precisely what is at issue in the cases we are considering. (...) But are those behaviors intentional in the relevant sense?" [Gendler (2008b), p. 565]. It is clear that (belief-discordant) alief-motivated behavior is not intentional *in the sense of being explainable by belief*. But is it intentional in a different sense? Gendler never explicitly says. Below, we introduce a wide notion of intentional explanation. Using that notion, we shall consider alief-motivated behavior as intentional action in a wide sense.
²⁴ For a more detailed discussion and defense of this picture of content, see Heck (2007) and Glüer (2013).
²⁵ Mandelbaum (2013) argues that a robust notion of alief faces a dilemma similar to ours: To do the explanatory work they are supposed to do, Mandelbaum argues, aliefs need to have propositional content. But if they have propositional content, it becomes unclear how they differ from already familiar psychological states.
²⁶ Imagining a bottle of poison or a cliff in front of one's feet is not sufficient for being motivated to act. Sure, sometimes a very lively imagination can make you jump back as if from a cliff in front of your feet. But if that happens, what was going on went beyond mere imagination. In such cases we say things like: 'Oh, my imagination was so real, so lively, that for a second it really seemed like there was this cliff.'

Or: 'For a second, it really was as if there was a cliff.' The state, that is, had become a state that not merely had a representational content, but one that represented the world as being as represented. If only for a second.

[27] Cf. Glüer (2009). See also Martin (2002), Pryor (2005), and Burge (2003) for what is in effect the same distinction. Burge helpfully calls strong representational states 'committal', adding the comment that it need not always be the 'whole subject' that commits to the truth of *p* – it might be a subsystem of the subject's cognitive architecture. He applies this idea to perceptual experiences in order to deal with phenomena such as known illusion. In effect, this amounts to a kind of 'compartmentalization' where commitments made by subsystems such as the perceptual system are kept separate from those of a subject's belief system. The idea seems eminently apt for capturing at least some of the alief-phenomena.

[28] The same identification of a 'content' activating the alief and the alief's content seems to underlie Gendler's remarks about alief-contexts being "hyperopaque". According to Gendler, alief-contexts are such that not even terms the synonymy of which the subject is fully aware of can be substituted salva veritate. This is because words or representations that the subject knows to be synonymous can nevertheless activate very different aliefs. Gendler's example is 'not poison' and 'safe to consume'. What alief-contexts presumably do, however, is report alief contents, not the stimuli activating them. Consequently, an alief activated by a sign reading 'not poison' is correctly reported by means of something like '*S* alieves that this is poison.' Moreover, such a report would seem to be exactly as (hyper-)intensional as any old belief-report.

[29] Even though one might wonder about the explanatory significance of the *B*-component now. If accepting the content motivates the behavior, adding activation of a motor routine to the 'content' of the state might seem redundant. Any mental state motivating behavior would – somehow – result in the activation of a motor routine, after all. In a way, our dilemma could have been formulated already in these terms, for the question can be turned around: If alief contains activation of a motor routine as part of its 'content', isn't that sufficient for explaining the ensuing behavior? What further contribution could the content make? Again, it might be that Gendler thinks of the content as activating the *A* and *B* components, somewhat in line with her already observed tendency to think of the content of the alief as its stimulus. In any case, there seems to be some tension here.

[30] The characteristic reason-providing role of belief for further belief is spelled out in some detail in Glüer (2009), – where it is further argued that perceptual experience is best construed as a (peculiar) kind of belief precisely because it plays that very role.

REFERENCES

BURGE, T. (2003), "Perceptual Entitlement", *Philosophy and Phenomenological Research*, LXVII, pp. 503-48.

BYKVIST, K. and HATTIANGADI, A. (2007), "Does Thought Imply Ought?", *Analysis* 67, pp. 277-85.

CURRIE, G. and RAVENSCROFT, I. (2002), *Recreative Minds*, New York, Oxford University Press.

DAVIDSON, D. (1963), "Actions, Reasons, and Causes", in: *Essays on Actions and Events*, Oxford, Clarendon Press, 1980, pp. 3-19.

–– (1971), "Agency", in: *Essays on Actions and Events*, Oxford, Clarendon Press 1980, pp. 43-62.

–– (1973), "Freedom to Act", in: *Essays on Actions and Events*, Oxford, Clarendon Press 1980, pp. 63-81.

–– (1987), "Knowing One's Own Mind", *Proceedings and Addresses of the American Philosophical Association* 60, pp. 441-458.

DOGGETT, T. and EGAN, A. (2007), "Wanting Things You Don't Want. The Case for an Imaginative Analogue of Desire", *Philosopher's Imprint* 7, pp. 1-17.

DRETSKE, F. (2000), "Norms, History, and the Constitution of the Mental", in*: Perception, Knowledge and Belief. Selected Essays*, Cambridge, Cambridge University Press.

FUNKHOUSER, E. and SPAULDING, S. (2009), "Imagination and Other Scripts", *Philosophical Studies* 143, pp. 291-314.

GENDLER, T. (2003), "On the Relation between Pretense and Belief", in: *Imagination and the Arts*, McIver Loopes, D. and Kieran, M. (eds.), New York, Routledge, pp. 125-142.

–– (2008a), "Alief and Belief", *Journal of Philosophy* 105, pp. 634-663.

–– (2008b), "Alief in Action (and Reaction)", *Mind and Language* 23, pp. 552-585.

GLÜER, K. (2009), "In Defense of a Doxastic Account of Experience", *Mind and Language* 24, pp. 297-373.

–– (2013), "Looks, Reasons, and Experiences", forthcoming in: *Does Perceptual Experience Have Content?*, Brogaard, B. (ed.), Oxford, Oxford University Press.

GLÜER, K. and WIKFORSS, Å (2009), "Against Content Normativity", *Mind* 118, pp. 31-70.

–– (2010), "The Truth Norm and Guidance: a Reply to Steglich-Petersen", *Mind* 119, pp. 757-761.

–– (2013), "Against Belief Normativity", forthcoming in *The Aim of Belief*, Chan, T. (ed.), Oxford, Oxford University Press.

HECK, R. G. (2007), "Are There Different Kinds of Content?", in: *Contemporary Debates in the Philosophy of Mind*, Cohen; J. and McLaughlin, B. (eds.), Oxford: Blackwell, pp. 117-138.

HUDDLESTON, A. (2012), "Naughty Beliefs", *Philosophical Studies* 160, pp. 209-222.

HÄGGQVIST, S. (2013), "Teleosemantics: Etiological Foundations"; *Philosophy Compass* 8, pp. 73-83.

MANDELBAUM, E. (2013), "Against Alief", in *Philosophical Studies* 165, pp. 197-211.

MARTIN, M. G. F. (2002), "The Transparency of Experience", *Mind & Language* 17, pp. 376-425.

MILLIKAN, R. (1993), *White Queen Psychology and Other Essays for Alice*, Cambridge, Mass., MIT Press.

NAGEL, J. (2012), "Gendler on Alief", *Analysis* 72, pp.774-788.

NICHOLS, S. and STICH, S. (2003), *Mindreading*, Oxford, Oxford University Press.

O'BRIEN, L. (2005), "Imagination And the Motivational View of Belief", *Analysis* 65, pp. 55-62.

PAPINEAU, D. (2013), "There Are No Norms of Belief"; forthcoming in *The Aim of Belief*, Chan, T. (ed.), Oxford, Oxford University Press.

PRYOR, J. (2005), "There Is Immediate Justification, in Steup, M. and Sosa, E. (eds),
       Contemporary Debates in Epistemology, Oxford, Blackwell, pp. 181-201.
SHAH, N. (2005), "How Truth Governs Belief", *Philosophical Review*, 112, pp. 447-82.
SHAH, N. and VELLEMAN, D. (2005), "Doxastic Deliberation", *Philosophical Review*
       114, pp. 497-534.
VELLEMAN, D. (2000), *The Possibility of Practical Reason*, Oxford, Oxford
       University Press.
WILLIAMS, B. (1973), "Deciding to Believe"; in: *Problems of the Self*, Cambridge,
       Cambridge University Press, pp. 136-51.
WILLIAMSON, T. (2000), *Knowledge and Its Limits*, Oxford, Oxford University Press.
ZALABARDO, J. L (2010) "Why Believe the Truth? Shah and Velleman on the Aim of
       Belief", *Philosophical Explorations* 13, pp. 1-21.