AIMING AT RATIONALITY: AN ALTERNATIVE ACCOUNT OF THE TRUTH-BELIEF RELATIONSHIP

Nikolai Viedge (31274)

A thesis submitted to the Faculty of Humanities, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Doctor of Philosophy.

Johannesburg, March 12, 2014

Abstract

One of the hot topics in doxastic epistemology at the moment is how to characterise the relationship between beliefs and truth. The extant literature is dominated by two views; a teleological understanding of the relationship – championed by people such as Asbjørn Steglich-Petersen – and a normative understanding of the relationship – championed by among others Nishi Shah. I argue that neither view does an adequate job of capturing the relationship between beliefs and truth. I argue that these two views should be abandoned in favour of a third model that sees beliefs as part of a doxastic system that aims at rationality.

Declaration

I declare that this dissertation is my own unaided work. It is submitted for the degree of Doctor of Philosophy in the University of the Witwatersrand. It has not been submitted before for any other degree or examination in any other university.

Candidate:
Nikolai Viedge

Signature: ……
Date: ……

For Jane and Liam Byrne.

Acknowledgements

First and foremost, I'd like to thank David Martens whose dedication, support and commitment has taken me from a vague idea about beliefs to the finished product. Any errors contained within are all my own and in no way represent the tireless work that David has put in to help me finish this project.

From start to finish there have been a number of people who have helped me take this project from an idea to the final product. Many of you I will forget to mention by name, but whose questions and comments at the various conferences where I've presented parts of this thesis, or conversations I've had with you over coffee/beer have helped me enormously either academically or emotionally over the last few years.

Of those who I can remember and would like to thank, I'd like to start by thanking the Rhodes University Philosophy Department *circa* 1998-2003. Without the patience and guidance of Francis Williamson, Marius Vermaak, Samantha Vice, Tom Martins and in particular Ward Jones I wouldn't have fallen in love with philosophy and I certainly wouldn't have had the confidence to attempt a PhD.

Outside of that Rhodes Philosophy Department, I'd like to thank Thad Metz, whose willingness to take the time to talk to me and provide invaluable suggestions helped me go from a PhD in theory to a PhD in practice.

For helping me to get the ball rolling and their unwavering support, I'd like to thank my parents, Ingo and Antje, and my siblings Erich, Myra and Frederick, who helped me push the ball over the line.

I'm particularly indebted to Jason van Niekerk for his constant words of advice, encouragement and support during this process and to Shahir Chundra and Brian Penrose whose independent words of support helped me get through a particularly rough week in terms of the academic side of my recent life.

I’m also enormously grateful to Daniel Whiting, Timothy Chan, Elisa Galgut, Michael Hicks and Andrew Reisner who were kind enough to share their recent work with me and for their encouraging words of support.

My thanks also go out to all the staff at St Augustine’s (in particular the Corridor Crowd) for their support and help, and to the Wits Philosophy Department who have helped see me through this journey.

To all my friends, of whom I have seen too little recently, I look forward to seeing you all again soon. Particular thanks go to Shahir, Shahan, Shameme and Jow’hara Chundra, Wendy Duncan, Yasaar Moosa, Shakti Satyapal, Anton and Luresha Chetty, Grant “G-Star” Sissons, Jono and Simon Ferreira, John Masterson, Chris Wareham and Bea Toniolo, James and Jade Vos, for giving me the time and the space to disappear without recrimination.

To the apple-juice-appreciation society, G and Rene, thanks lads I’ve needed to appreciate the finer points of apple juice for some time and I’m grateful for your help in this area. To Paddy and J, the boys’ nights out have been very necessary and greatly appreciated. I’m really looking forward to the next one.

To Richard “Ricky F” Flockemann thank you for many helpful chats about doxastic normativity in which you’ve kindly refrained from telling me that I don’t know what I’m talking about.

Big thanks also go out to the VCPs for their humour, support, encouragement and help. In particular, thank you to Charles and Margaret, who were kind enough to turn their home into my office away from home.

Most importantly, I’d like to thank my wife Jane who has been incredible in so many, many ways. I can never repay the faith, support, kindness and understanding you’ve shown me during this time. Thank you so much.

Finally, I am indebted to financial support I received, at different times, from the National Research Foundation and the University of the Witwatersrand during the process of writing this dissertation.

Contents

Title page	i
Abstract	ii
Declaration	iii
Dedication	iv
Acknowledgements	v
Contents page	viii
Chapter 1: Introduction	1
Chapter 2: Beliefs	6
Chapter 3: The teleological account	68
Chapter 4: The normative account	119
Chapter 5: Aiming at rationality	160
Chapter 6: Conclusion	219
Bibliography	225

Chapter 1: Introduction

There is an argument hidden in the literature around the truth-belief relationship that it is the job of the truth-belief relationship to explain the features of belief.

David Owens (2003), for example, argues that what makes a teleological account of the truth-belief relationship compelling is the explanatory force of this model in terms of its ability to explain the features of beliefs.

Nishi Shah (2003), meanwhile, argues that teleological accounts of the truth-belief relationship should be abandoned in favour of a normative understanding of the truth-belief relationship in virtue of the fact that the former can't explain, what he calls, transparency, while the latter can.

In both instances, the claim that both are alluding to quite clearly, without coming out and saying it outright, is that a proper account of the truth-belief relationship needs to be able to explain the features of belief in virtue of that relationship.

While there is plenty of literature about the relationship between beliefs and truth – which spawns from Bernard Williams' (1973) "Deciding to Believe" paper – as well as literature that speaks about various features of belief, either directly or obliquely, none of these projects to my knowledge a) explicitly makes the claim that the job of the truth-belief relationship is to explain the features of belief and/or b) looks at several features of belief and examines the existing accounts' ability to explain them.[1]

The aim of this project is to do just that. I begin by laying out several features of belief before examining the ability of the two major accounts of the truth-belief relationship (teleological accounts and normative accounts) to explain these features.

The existing literature around the truth-belief relationship appears to have focussed on several features of belief: either showing how their account can make sense of these

[1] Among those who have talked about one or more features of belief see Andrew Huddleston (2012), Conor McHugh (2012), Shah (2002:2003), Shah and David Velleman (2005), Velleman (2000), Bernard Williams (1973:2002), Ward Jones (2002:2004), David Owens (2002:2003), Dion Scott-Kakures (2000) and Asbjørn Steglich-Petersen (2006:2009).

features or criticising another account for not being able to explain the feature(s) in question. Some of those features regularly highlighted in the literature are that: beliefs regulate for truth; only epistemic reasons can be seen by the agent as reasons for his/her belief; and some of our beliefs are based, unbeknownst to us, on non-epistemic reasons (motivated beliefs).

However, this list seems woefully incomplete. In addition to the above features of belief is at least one feature of belief that I believe is in need of explanation, namely the fact that some of our beliefs display a remarkable level of evidence resistance.

I have been fascinated by the fact that some beliefs are resistant to evidence from the moment I read and understood David Hume's (1977) argument against the reliability of induction.

That these evidence-resistant beliefs (*idées fixe*) exist and that a sliding scale of evidence resistance is a feature of beliefs – from beliefs that regulate immediately in the face of counter evidence to, seemingly, cannot be shifted – have struck me as obvious truths about our doxastic lives.

The continued existence of bigoted beliefs in the face of overwhelming evidence to contrary of the truth of those beliefs has struck me as a particularly poignant example of both – particularly in light of new laws in Nigeria and Uganda both outlawing and allowing for the persecution of homosexuality and homosexuals.

However, despite this being a commonplace feature of belief the extant literature does nothing to address how it is possible to have beliefs that maintain in the face of, sometimes compelling, counter evidence. Velleman (2000) comes the closest to exploring these beliefs when he looks at delusions, but argues that because delusions don't respond to counter evidence they ought not to be considered as beliefs.

Having carefully gone through the literature and thought about the existing models of the truth-belief relationship I've realised that not only have neither of these models thought about how to account for this feature of belief, neither of them can explain how it is that some beliefs maintain in the face of counter evidence.

As a result, in what follows I suggest a new model of the truth-belief relationship that I believe can, among other things and for the most part, provide a coherent explanation for why beliefs admit to a sliding scale of evidence resistance.

In terms of methodology my starting point is the view put forth by Williams (1973). As the majority of commentators I deal with are either responding directly or indirectly to Williams' project, discussions about the truth-belief relationship as well as the doxastic voluntarism debate from this viewpoint are appropriate.

Williams (1973) notes that his is a discussion of beliefs "as a psychological state … [of] people believing things" (Williams, 1973:136). Thus, my project is not an attempt to define or defend some ontology of belief, instead it is interested in working out what we can say of our beliefs by examining those beliefs from an ordinary, everyday, experiential perspective.[2]

It is also worth noting that there are several issues which I'm skirting in this dissertation. While a lot has been written about beliefs aiming at truth, there is another school of thought that suggests that beliefs aim at knowledge rather than at truth. While this is certainly a worthwhile understanding of how to understand the relationship between belief and truth, it is one that I'm avoiding.[3] I feel that by showing that there are other ways of cashing out the truth-belief relationship – by suggesting an alternative account that has neither truth, nor knowledge, as the target – I'm offering something worthwhile to the literature.

I am also dodging the issue of degrees of belief. Eric Schwitzgebel (2010) notes that we may have more confidence in some beliefs than we do in others, such that there is a degree of belief. As he puts it:

> One common way of formalizing this idea is by means of a scale from 0 to 1, where 0 indicates absolute certainty in the falsity of a

[2] For an account of the ontological theories of belief see Eric Schwitzgebel (2010)
[3] For a critical overview of the existing arguments for and against the claim that beliefs aim at knowledge, see Daniel Whiting (2013).

> proposition, 1 indicates absolute certainty in its truth, and .5 indicates that the subject regards the proposition just as likely to be true as false. This number then indicates one's "degree of belief". (Schwitzgebel, 2010)

To my mind I can't see how to make sense of this. It seems to me, given what beliefs are, either we believe something or we don't. To say of something "I believe that *p*, but I'm not certain of it" suggests one of two things. One, I used to believe it, but now, under consideration I suspend my judgement until it a) creeps back in, b) I confirm it or c) I discredit it. Or, two, I still believe it, but the reasons I have may not convince you.

While there may well be a genuine doxastic phenomenon here, I don't think I need to engage in it in order to achieve the goals of my project. What I have to say is about the relationship between beliefs and truth and should apply even if we do have degrees of belief.

In what follows in Chapter 2 I either lay our or argue for several features of belief that I think an account of the truth-belief relationship must be able to make sense of. Having done so, I make explicit the argument above about how accounts of the truth-belief relationship need to make sense of the features of belief. As I've indicated, this is not a new argument, but it is one that needs to be stated baldly.

In Chapters 3 and 4 I examine the ability of teleological and normative accounts of the truth-belief relationship respectively, and examine their ability to make sense of the features of belief.

Having discredited these accounts' ability to make sense of several features of belief I then put forward my own account in Chapter 5.

I argue in favour of two claims. The first is that we shouldn't see beliefs as having a relationship with truth other than the conceptual one. The second is that we should see beliefs as part of a doxastic system that has the goal of truth, tempered by other cognitions, conations and emotions.

These tempering factors, I argue, mean that the doxastic system can be said to *de facto* aim at rationality.

I argue that what stands in favour of this rather complicated model of belief is that it explains the majority of features of belief.

Chapter 2: Beliefs

1. Introduction

In Bernard Williams' (1973) paper "Deciding to Believe" are four words that have been only marginally less responsible for the creation of paper than the Harry Potter novels: "beliefs aim at truth" (Williams, 1973:136).

These four words are the crux of Williams' (1973) very short account of how to make sense of the relationship between beliefs and truth, and as the subsequent proliferation of literature suggests it's open to interpretation.[4] This literature has tried to determine the correct way to understand Williams' metaphor, trying to understand the relationship between belief and truth.

In the extant literature there are two main ways that this phrase has been cashed out: those who have taken Williams more-or-less literally, arguing that we should see the relationship as teleological, and those who think the phrase should be understood normatively.[5]

I, in my own attempt to aid in the demise of paper-producing trees, disagree with both. I argue that both of these theories admit to damning objections in the existing literature, and that neither account is able to capture the features of belief. I argue that rather than trying to try to paper over the cracks of one to meet the objections of the other we should abandon both.

In brief, my argument against the prevailing models is as follows. Any adequate account of the truth-belief relationship must be able to, in virtue of that relationship, explain the features of belief. A failure to explain the features of belief undermines the account's characterisation of the truth-belief relationship. The two main accounts of the truth-belief relationship prevalent in the literature – in addition to having problems

[4] As Timothy Chan (2013) puts it, "Williams's explication [of what aiming at truth means] in his article is brief, amounting to little more than two paragraphs (1973:136-7). This is because belief aiming at truth is in Williams's paper an explanans, with the explanandum being the impossibility of believing at will" (Chan, 2013:1).

[5] Chan (2013) describes these two views of the truth-belief relationship as "the primary … fault line in the debates about the aim of belief" (Chan, 2013:3).

of their own – also fail to capture the features of belief. Because of this neither of these accounts should be considered as the definitive account of the truth-belief relationship. Therefore, a new account of the truth-belief relationship is needed.[6]

Each of these claims needs to be explained and defended and in what follows I aim to do just that.

In this chapter I begin by exploring a number of features of belief. Having done so, I make explicit an argument in the literature that it is the job of the truth-belief relationship to be able to explain the features of belief.

Chapters 3 and 4, respectively, examine teleological accounts and normative accounts, detailing them and examining their suitability as accounts of the truth-belief relationship. I argue that both of these accounts of the truth-belief relationship fail.

In Chapter 5 I present my own account. I argue that my account is better able to make sense of the features of belief than the two main rival theories. I also consider some objections to my account

The structure of this chapter is as follows.

In Section 2 I lay out what is meant by the term "belief".

In Section 3 I examine several important features of belief. In addition, I explain what I think is meant by "truth" in the literature. The literature on how to understand Williams' (1973) "aim at truth" (1973:136) metaphor deals more or less exclusively

[6] Shortly before submitting my dissertation I became aware of an Honours thesis by Anna Edmonds (2007) whose argument has a similar structure to mine. She argues that the claim that beliefs aim at truth is false. She claims since our beliefs routinely admit to non-epistemic reasons that neither we as agents, nor beliefs themselves can be said to aim at truth. In addition, she argues that a normative understanding of this claim could not force our beliefs to act in line with truth considerations. She argues that the goal of cognition is instrumental usefulness. While the structure of her paper is similar, there is very little overlap in terms of actual argument. My argument rests on how to make sense of this claim as a statement about the truth-belief relationship. In what follows, I argue that the notion that beliefs aim at truth is essentially correct. In those cases where there is overlap between my ideas and Edmonds, I will highlight the similarities and differences. In all cases while Edmonds' project predates mine, my arguments and ideas were arrived at independently of her thesis.

with the truth-belief relationship, with next to nothing being said about what is meant by truth in this context. Part of Section 3 looks to address this gap.

In Section 4 I highlight an argument in the literature that any account of the truth-belief relationship must be able to account for the features of belief.

2. What is a belief?

A full explanation of what a belief is, and the ability to distinguish beliefs from other similar propositional mental states, such as assumptions and guesses, requires us to have an understanding of the relevant features of belief.

However, understanding why these features are linked with this mental state requires an idea of what type of mental attitude is picked out by "belief", helping us understand the features of belief as well as what we mean by truth.

Eric Schwitzgebel (2010) notes, "Contemporary analytic philosophers of mind generally use the term 'belief' to refer to the attitude we have, roughly, whenever we take something to be the case or regard it as true" (Schwitzgebel, 2010).

However, while Schwitzgebel's (2010) characterisation of belief is generally accepted the problem is that as it stands it is too broad to clearly differentiate beliefs from other similar mental propositional attitudes.

As Nishi Shah and David Velleman (2005) point out: "Believing that *p*, assuming that *p*, and imagining that *p* [all] involve regarding *p* as true – or, as we shall call it, accepting *p*" (Shah & Velleman, 2005:497).

Assuming that *p* is to take it that *p* is true. In assuming that this example is clear I am taking it as true that this example is clear. However, this in no way suggests that I believe that this example is clear. There are clear differences between the two such that when I believe that *p* I take *p* to be true in a way that is fundamentally distinct from when I assume *p* to be true. And the same can be said for guessing that *p* or imagining that *p*.

So, "[w]hat distinguishes belief from the other modes of acceptance?" (*ibid*.).

There is a fair bit of literature around this. Andrei Buckareff (2010) notes that there are at least six ways that belief and acceptance can be differentiated:

> Many would claim that belief and non-doxastic acceptance differ in some or all of the following six ways. First, belief aims at truth, while acceptance aims at utility or success. Second, belief is shaped by evidence; acceptance need not be shaped by evidence. Third, belief is context-independent insofar as it is not shaped by an agent's purposes, but acceptance is often context-dependent and shaped by an agent's purposes. Fourth, belief is subject to an ideal of agglomeration, and acceptance is not regulated by any such ideal. Fifth, belief comes in degrees, while acceptance is all or nothing. Finally, belief is not subject to direct voluntary control, while acceptance can be under our direct voluntary control ... (Buckareff, 2010:255)

There is a seventh option, one that Shah and Velleman (2005) propose, namely

> … that conceiving of an attitude as a belief, rather than an assumption or an instance of imagining, entails conceiving of it as an acceptance that is regulated for truth, while also applying to it the standard of being correct if and only if it is true. (Shah & Velleman, 2005: 497)

A number of these are contentious. That beliefs are involuntary, for example, is the subject of vigorous debate and, thus, seems like a dangerous condition upon which to hang the distinction.

Trying to differentiate beliefs in virtue of a norm of correctness is equally tricky. David Owens (2003) notes that truth is not only the standard of correctness for belief it is also the standard of correctness for guesses; a guess is correct if, and only if, it is

true (Owens, 2003:290). Thus, trying to use this norm as a differentiating feature is both contentious and requires a fair bit of work.

The concern is legitimate. There is a real difference between an assumption and an imagining and a belief. All three spell out completely different propositional attitudes and the debate about how to separate belief from other types of acceptance is lengthy and ongoing.[7]

I will not say a lot about this debate, nor discuss the merits of the various approaches. The debate is too complicated to do justice to it here. Highlighting deficiencies in each of these approaches in order to argue for what I take to be the right way of distinguishing between beliefs and other similar mental propositional attitudes is a thesis on its own.

What is required for my purposes is a description of the concept of belief strict enough to be able to differentiate beliefs from other forms of acceptance, without losing the intuition behind Schwitzgebel's (2010) idea above.

Hamid Vahid's (2006) description, I feel, best captures what it is to be a belief.

> … while assuming or imagining that p involve regarding p as true in order to reach certain specific goals, believing p involves no such

[7] Andrei Buckareff (2006) in his first footnote offers an overview of some of the literature surrounding this debate.

> For recent treatments of the differences between belief and acceptance see the following: William P. Alston, 'Belief, Acceptance, and Religious Faith', in Jeff Jordan and Daniel Howard-Snyder (eds) *Faith, Freedom, and Rationality: Philosophy of Religion Today* (Lanham: Rowman & Littlefield, 1996), pp. 3–27; Robert Audi, 'Belief, Faith, and Acceptance', *International Journal for Philosophy of Religion* , 63 (2008), pp. 87–102; Michael Bratman, 'Practical Reasoning and Acceptance in a Context', *Mind* , 101 (1992) pp. 1–15, especially 2–9; Andrei A. Buckareff, 'Acceptance and Deciding to Believe', *Journal of Philosophical Research* , 29 (2004), pp. 173–90, and 'Can Faith be a Doxastic Venture?', *Religious Studies* , 41 (2005), pp. 435–45; L. Jonanthan Cohen, *An Essay on Belief and Acceptance* (Oxford: Clarendon, 1992); Pascal Engel, 'Believing, Holding True, and Accepting', *Philosophical Explorations*, 1 (1998), pp. 140–51; Robert C. Stalnaker, *Inquiry* (Cambridge, MA: MIT Press, 1987), Ch. 5; Raimo Tuomela, 'Belief versus Acceptance', *Philosophical Explorations*, 3 (2000); 122–37, Bas C. Van Fraassen, *The Scientific Image* (New York: Clarendon, 1980); and J. David Velleman, 'On the Aim of Belief', in his *The Possibility of Practical Reason* (New York: Oxford University Press, 2000), pp. 244–81. While each of these authors defends the claim that there is a difference between belief and acceptance, there are important differences between each. (Buckareff, 2006:259-260)

> distinct goals. To believe p is to regard p as true for its own sake, not to regard p as true "for the sake of something else". In other words, regarding-p-as true is not a *means* to some exterior end but an end in itself. (Vahid, 2006:324)[8]

When I 'assume' that my example is clear I assume that it is clear to be able to move on with the rest of this chapter. When I 'imagine' that I have completed my project I do so to try to experience what that might feel like. When I 'guess' that England will win the 2013 Ashes cricket series in England 2-1, I do so in order to win a bet. However, when I 'believe' that I'm struggling for clarity I do not believe it for the sake of anything. I just take it as true. Thus, to believe that *p* is to take *p* as actually being true for no extrinsic reason, but simply because one takes *p* as being true, *tout court*.

This is not say that beliefs convey truth onto their propositions: that in some way by believing *p*, *p* is somehow made true. The claim here is what we take as true is a belief, rather than beliefs creating truths.

As Lloyd Humberstone (1992) highlights, while there is a lot of debate about how best to cash out the direction of fit claim (the question about whether truth causes belief or belief causes truth), it is accepted that direction of fit for belief is from truth to belief rather than from belief to truth (Humberstone, 1992:59-61).[9]

In addition, this notion of belief doesn't commit us to only having true beliefs. While I say more about truth in a doxastic context later, the important thing to remember is

[8] There is one caveat to this claim that needs to be added. In cases of motivated believing the agent forms the belief that *p* precisely because of a reason that is extrinsic to the fact that he/she takes *p* as being true. The person who believes that his/her son is a visionary, rather than delusional, despite evidence to the contrary forms this belief as a result of reasons that are extrinsic to the truth of the matter. However, as I argue below, this reason is opaque to the believer. He/She cannot see that his/her belief is based on a non-epistemic reason. Hamid Vahid's (2006) claim, therefore, should be understood as to believe *p* is to take *p* as being true for no extrinsic reason *that is transparent to the agent* other than the fact that he/she takes *p* as being true.

[9] It is worth noting that Kwasi Wiredu (1980) argues that truth just is what you believe (Wiredu, 1980:111-123). But this is a notion of what truth means, not of belief. If Wiredu is right, then all it means is that there can be no further adjudication in terms of correctness when it comes to two contradictory beliefs. However, this does not entail that your belief conveys truth onto the world. Wiredu notes that beliefs come from the evidence presented; highlighting that what informs your belief will be what you take as true, rather than your belief creating or conferring truth onto the propositions believed (Wiredu, 1980:117).

that truth as it pertains to beliefs is inclusive, rather than prescriptive. While it might well be the case that we ought only to believe true things, all this notion of belief commits us to is that if you think something is true, then we call that propositional mental attitude a belief.

With this notion of belief in mind, let's turn our attention to the features of belief.

3. Features of belief

In this section I examine several features of belief. While this list is by no means exhaustive I present those features that have been used in the literature to show why one account of the truth-belief relationship is a good one, because it can explain the feature in question, or a bad one, because it can't. In addition, I present a case for those features that I feel accounts of the truth-belief relationship should make sense of.

Some of the features below I simply stipulate and explain rather than argue for. In these cases the feature is non-contentious, or, given that this is philosophy, sufficiently non-contentious and intuitive enough for me to feel happy about claiming its existence. For those features that I feel are sufficiently contentious to merit further discussion I present an argument for my point of view.

I begin with one feature of belief that in conjunction with what it means to be a belief dictates and helps us understand several other features of beliefs: namely that the relationship between belief and truth is direct (3.1). In section 3.1 I also say something about what truth means in this project and, I take it, in the literature more generally (3.1.1). This discussion of truth and the discussion in 3.1 helps us understand the scope of beliefs (3.2). In section 3.3 I note that we have dispositional as well as occurrent beliefs. In section 3.4 I highlight that the only reasons that count as reasons for belief are (what an agent takes as) epistemic reasons.

I then look at Shah's (2003) transparency claim (3.5). Shah and later Shah and Velleman (2005) claim that transparency is a feature of belief that appears only in doxastic deliberation. I argue for a modified version of this claim, claiming that if

transparency is a feature of belief it must belong to all beliefs, not just deliberative beliefs. I argue in section 3.5.1 that what projects like Pascal's Wager and doxastic deliberation reveal about beliefs must pertain to beliefs, *simpliciter*.

In section 3.6 I note that what suggests a strong link between belief and truth is the fact that beliefs regulate for truth. This regulation for truth along with the direct relationship between belief and truth has been thought by many (following reasoning famously explicated by Williams (1973)) to show that beliefs are beyond our control. In section 3.7 I look at the case for doxastic involuntarism. I argue while the case for the conceptual argument for doxastic involuntarism may not go through, it certainly seems impossible for an agent to believe at will.

I then examine wishful thinking (3.8), biased beliefs (3.9), false beliefs (3.10), beliefs based on inconclusive evidence (3.11) and evidence resistant beliefs, *idées fixe* (3.12). The reason I talk about any kind of wonky belief (3.8, 3.9, 3.10 and 3.12) is not only because they are fascinating, but because they, in particular *idées fixe*, have been given scant attention in the literature in terms of the truth-belief relationship.

3.1. The direct relationship between belief and truth

The first thing to note about the relationship between truth and belief is that it is direct.

As Shah (2003) explains:

> Truth is not an optional end for first-personal doxastic deliberation, providing an *instrumental* or *extrinsic* reason that an agent may take or leave at will. Otherwise there would be an inferential step between discovering the truth with respect to *p* and determining *whether to believe that p*, involving a bridging premise that it is good (in whichever sense of good one likes, moral, prudential, aesthetic, all-things-considered, etc.) to believe the truth with respect to *p*. But there is no such gap … The seamless shift … is not a quirky feature

> of human psychology, but something that is demanded by the nature of first-person doxastic deliberation. (Shah, 2003:447)

While Shah (2003) is talking only about deliberative beliefs here, I think it is safe to take this to be a feature of beliefs generally speaking.[10] As Pamela Hieronymi (2006) notes: "When you change your mind about what is true, you have therein, *ipso facto*, changed your beliefs" (Hieronymi, 2006:53). This is true of more than just deliberative beliefs. When you think that something is true you believe it, irrespective of whether or not you're deliberating about it. And you do so without any inferential step which involves "a bridging premise that it is good … to believe the truth with respect to *p*" (Shah, 2003:447).

It is as a result of this direct link between truth and belief that your perceptual beliefs, for example, tend to change the moment your perception changes. If I have a belief that there is a glass on the table, the moment I see the glass being taken off the table I no longer believe that it's on the table.[11]

In addition, no matter how much you might not want to have a belief, the moment you take something as true for no extrinsic reason that's transparent to you other than you think it's true you have the belief.

There's a scene in the movie *The Hangover* where one of the characters, Alan, after a heavy night out, goes to the bathroom. Standing at the toilet, urinating, Alan glances to the right where unexpectedly there happens to be a large tiger. While Alan would undoubtedly prefer not to be in a bathroom with a large and scary looking tiger, no practical reason, no matter how beneficial, interferes with him forming a belief that there is a tiger in the bathroom with him.

---

[10] Technically, the quote above is part of Nishi Shah's (2003) defence of transparency. While I think that transparency and the direct relationship between beliefs and truth are subtly distinct I do think that this quote from Shah captures rather neatly what I'm terming the direct relationship between beliefs and truth. I say more about this distinction later on in my section on transparency. If, however, transparency just is the direct link between beliefs and truth then the quote is obviously applicable. It is also worth noting that this notion of the direct relationship between beliefs and truth that Shah provides matches rather neatly onto Vahid's (2006) notion above of what a belief is.

[11] There are of course exceptions. If one is going to a magic show then one might quite literally not believe what your eyes are telling you. However, in ordinary cases of perception I take it that we believe in line with our perceptions. I'm grateful to David Martens for highlighting this point.

In this case, where Alan is just about incapable of deliberating, the link between what he takes to be true and what he believes is direct, it is not an optional extra.[12]

It is important to note that there is a distinction here between what you or I or Alan takes as truth, which informs the belief, and the actual truth; if indeed something like objective truth exists. The former is what has a direct link with belief. It is what we take as true that is the thing that has a direct link with what we believe. While the latter may coincide with what you or I or Alan take as true, this isn't a condition that needs to be satisfied in order for a belief to obtain.

This distinction and, as a result, what is meant by truth in this debate is important and so I turn now to what I think is meant by truth in this area.

3.1.1. Truth

The claim that there is a direct link between belief and truth is, I take it, fairly uncontroversial.

When I see that there is a glass on the table, I believe that there is a glass on the table and when someone removes the glass from the table I no longer believe that there is a glass on the table, *ceteris paribus*.

However, it may be the case that something like René Descartes' (1996) external world scepticism is correct that there is no external world and so there is no glass. In

[12] Simon Beck, in conversation, has questioned whether or not Alan has a belief. Beck thinks that Alan could be seen as a quintessential example of someone in the grips of an information state. (An information state would be a sort of mental reflex that while motivating certain behaviour doesn't have a truth value in the same way that a belief does and as a result fails to qualify as a belief). Given that we have unconscious beliefs (I defend this is section 3.3 below) then the only reason that Alan could be said to have something other than a belief would be if Alan didn't think it was true that there was a tiger in the bathroom with him. In this case perhaps something like the following happens in Alan's brain: "Danger! Get out of here!" If this is what has happened then it appears that Alan has an information state, rather than a belief. Once Alan has left the bathroom he might then reflect on why he left the bathroom and come to the belief that he left because there was a tiger. I am quite happy to accept that Alan's behaviour can be as easily described by an information state as it can by a belief. However, I take it that Alan's running out of the bathroom can be just as easily explained by a pre-reflective belief that there is a tiger in the bathroom as it can by Alan having an information state of danger. The point that is being made here is that the moment one takes something as true there is a direct link between this and forming a belief and that this process needn't be a conscious, reflective mental process.

this case, if there is a direct link between beliefs and some objective notion of truth I ought to have no beliefs about glasses whatsoever.

Further, as noted above, simply believing that *p* doesn't mean that *p* is true. All it means is that you take *p* to be true. You *think p* is true. But, as you no doubt have discovered, we get things wrong depressingly often. Believing that *p* doesn't mean getting it right.

So what exactly do we mean by true or truth when we say that beliefs have a direct link with truth? What exactly do we mean by "truth" in the claim that "beliefs aim at truth" (Williams, 1973:136)?

While the literature of what the relationship is between truth and belief is ubiquitous almost nothing is said about what truth is in this literature. Williams (1973), Shah (2003), Velleman (2000), Shah and Velleman (2005), Owens (2003), and Vahid (2006), along with many others all talk about the relationship between truth and belief without saying what they mean by truth.

Discussions about truth come from a wide variety of angles and quite quickly become complicated. As Michael Glanzberg (2009) notes: "Truth is one of the central subjects in philosophy. It is also one of the largest … [and] has been a topic of discussion in its own right for thousands of years" (Glanzberg, 2009). Any attempt to define what truth is (for example) quickly becomes a project on its own.[13]

Furthermore, given that the discussions around beliefs and truth that I am interested in come from a first-person, experiential perspective, the need to say exactly what truth is from, say, a metaphysical perspective is beyond the scope of this project specifically and the literature more generally speaking.

The reason for this is that while metaphysical perspectives of truth, for example, are contentious and up for debate, what isn't contentious is that anyone who is reading

[13] For an overview on discussions around truth see Michael Glanzberg (2009).

this has beliefs. Complicating the debate by trying to resolve what is meant by truth metaphysically speaking is more problematic than useful in this context.

However, as the questions above indicate, this does not obviate the need to say something about truth as it pertains to belief and so I briefly sketch out what I mean by truth.

An account of truth as it pertains to belief must be able to account for the fact that people can quite legitimately have contradictory beliefs and can believe almost anything. Kwasi Wiredu (1980), for example, has anti-realist beliefs about truth, while Peter Van Inwagen (2009:93-108) has realist beliefs about truth. Both seem to be candidates to legitimately be called beliefs and so an account of truth as it pertains to belief must explain how it is possible that both can have beliefs, given that one of their beliefs must be false. It seems we have a paradox. Both Wiredu and Van Inwagen have beliefs that have a direct link with truth, but one of their beliefs must be false. How is this possible?

What is going on is that there are two different meanings of truth at play. Objective truth (Truth hereafter), whatever that might be, and what-we-take-to-be-Truth (hereafter, truth).[14]

What we believe is truth.

Now, remember, what we believe (truth) is what we at the time take as Truth. However, this doesn't mean that is actually is True, only that we think it's True. Imagine that you are shown a clever illusion of a bottle on a table. You are asked "do you believe there's a bottle on the table?". Assuming that the asking of the question doesn't raise your suspicions and assuming that you are unaware of the illusion and are not in a particularly sceptical frame of mind, you are likely to respond that "yes, you believe that there's a bottle on the table". At the time of responding, or at the time of holding the belief, you take it as True that there's a bottle on the table. However, it's not actually True; you just think it's True, your belief is true.

[14] In cases where I've yet to establish whether it is Truth or truth that's being talked about, I use the lowercase truth.

The direct link I've been talking about is between truth and belief, rather than Truth and belief. If there were a direct link between Truth and belief, then all our beliefs would be objectively True. Since this is patently false the direct link must be a direct link with truth.

In Plato's (2002) Allegory of the Cave, he remarks on those trapped in the cave: "To them, I said, the truth would be literally nothing but the shadows of the images" (Plato, 2002:374). Plato doesn't think that the shadows on the wall are the Truth of the matter, but he understands completely why those trapped in the cave would believe it. For those trapped in the cave the only truth they have access to is "the shadows of the images" (*ibid*.). They would take these shadows as the Truth and so form beliefs based on this truth, rather than the actual Truth of the matter.

This distinction explains why we encounter such varied beliefs. My truth needn't be your truth and getting to the Truth isn't always easy. As Daniel Whiting (2013) notes, what is True is "a matter which is, so to speak, out of one's hands ..." (Whiting, 2013:185). In the quest to find the Truth, even if that Truth is that there is no objective Truth, people are reliant on what is, or could be, or has been thought to be Truth.

Let's leave discussions of truth and Truth. For the moment Vahid's (2006) concept of belief, along with the distinction between truth and Truth is sufficient to help us understand what we mean by the direct link between beliefs and truth.

3.2. The scope of belief

This concept of belief as it pertains to truth means that the scope of beliefs is limited to any propositional claim that we *can* actually take as True. Thus, I am able to believe that *p*, if it is possible to think that *p* is True.

With only this very limited restriction in place, beliefs range from the profound (God does (not) exist) to the profoundly boring (I don't have clean underwear). Furthermore, what I believe needn't be what you believe. You may well believe you have clean underwear.

As a result, *p* is only literally unbelievable when one cannot believe the truth of the proposition in question. I couldn't, for example, believe that I am both awake and typing this at my office and at the same time sleeping in my bed. In this case the proposition cannot be true and so cannot be believed.[15]

If, however, you are uncertain of the truth of *p*, then you won't have a belief that *p* or not-*p*. In this case there is no proposition that you take to be True. Thus, if you happen to be unsure of, for example, whether or not other intelligent life forms exist in the universe, then you have no belief either way.

Given this very minimal restriction it follows that if you think that you actually are Napoleon, then you believe that you are Napoleon. I might think you're mentally ill, but while you continue to take it as True (in the way detailed above) that you are Napoleon you have the belief "I am Napoleon".

However, not everyone thinks that the scope of belief is so broad. Donald Davidson (1986) and Velleman (2000) both argue that what counts as a belief ought to be limited in certain ways.

3.2.1. Beliefs and the correct cause of belief

Davidson (1986) argues that in order for a mental propositional state that an agent takes to be True (hereafter MPS=T) to be called a belief, that MPS=T must be caused in the agent in the right way. Thus, if S has a belief that *p*, then the MPS=T that *p* was caused in S by *p*. If you take some proposition as True, but this mental state hasn't been caused in the right way then, according to Davidson, you have an MPS=T, but not a belief. As Davidson puts it: "…we must, in the plainest and methodologically most basic cases, take the objects of a belief to be the causes of that belief" (Davidson, 1986:317-318).

[15] I take it that it cannot be True either that one thing is in two places at the same time. However, perhaps this is possible and not being able to believe that something is in two places at the same time is merely a psychological limitation specific to humans.

Let's go back to the person who has a delusion that he is Napoleon. Davidson (1986) argues that since the MPS=T hasn't been caused in the right way – by actually being Napoleon – what the agent has is not a belief, but an MPS=T. Thus, the scope of beliefs should be limited to those MPS=Ts that meet this causal condition. However, while Davidson certainly raises an interesting point, I don't think his condition applies here.

The first thing to note is that Davidson (1986) is making a metaphysical point. That, metaphysically speaking, in order for an MPS=T to *really* be a belief then that MPS=T must be caused in the right way. From the perspective of this project, an experiential perspective, an MPS=T just is a belief. When I take *p* as being True for no extrinsic reason that's transparent to me other than I take *p* as True, this is a belief. This just is the concept of belief. The brain in the vat's MPS=T 'the sky is blue' (assuming a perfectly deceitful experiment) is indistinguishable from the *belief* 'the sky is blue'.

Of course, when we ask, "but what is the brain's MPS=T *really*?" we might, if we agree with Davidson (1986), claim that the "brain actually has no belief …" (Vahid, 2006:330-331). In this case the ultimate reality of the MPS=T in question is being appealed to. And, in the case of the brain in the vat, the MPS=T fails to count as a belief because the MPS=T has not been caused by a blue sky.

Countering Davidson's (1986) challenge by claiming that I'm not making a metaphysical claim is only partially helpful. Even though Davidson's project is a metaphysical one, this doesn't mean that his objection doesn't carry weight here.

Consider the following. From an experiential perspective there being a glass on the table (*p*) is the correct cause of my belief that *p*, if *p*, and my perception of *p* caused my belief that *p*. In this case, my MPS=T counts as a belief. However, imagine that I have been hit on the head by a cricket bat and as a result of this blow form an MPS=T that *p*. Here, without appealing to brains in vats or anything else, we might want to say that what I have is an MPS=T, but not a belief, because it was caused in the wrong way.

Ruling out those MPS=Ts that aren't caused in the right way as beliefs is potentially useful. For one, it means that an account of the truth-belief relationship needn't account for MPS=Ts, only beliefs. What this means is that motivated beliefs, delusional beliefs and perhaps even *idées fixe*, which are unlikely to meet Davidson's (1986) requirement, no longer have to be accounted for as they are no longer beliefs.

However, despite the potential benefits, Davidson's (1986) requirement would be so stringent a condition that to employ it from an experiential perspective would rule out any talk of beliefs at all.

Let's consider for a moment what it would be like to take this condition on board. Being able to adjudicate whether the MPS=T in question qualifies as a belief means being able to accurately tell an etiological story of that MPS=T: either in favour of that MPS=T being a belief or against it.

The first thing to note is that determining what counts as a belief would require an external adjudicator. From a first-person perspective a belief and an MPS=T are indistinguishable. In addition, as we shall see in more detail later on, one of the features of belief is that when we are asked about why we believe something we tell, as a matter of course, an epistemically rationalising story about that belief. Telling an epistemically rationalising story will also be telling, it seems to me, a story that highlights the belief being caused in the right way.[16] Further, it seems to me that the moment I think my belief hasn't been caused in the right way it disappears and so would my MPS=T. This is not to say that when we have a belief it *is* caused in the right way, just that it seems we must think it is, or stop having the belief.

As a result, we can't simply ask someone how his/her belief was caused and then determine if his/her MPS=T counts as a belief. In order to determine if what you have is a belief or an MPS=T we need an external adjudicator.

[16] Two points to make here. The first is that I'm assuming that reliable testimony from someone whose MPS=T was caused in the right way would allow the person with the MPS=T based on that testimony to have an MPS=T that was caused in the right way to constitute a belief. The second, as we shall see below, is that there will be exceptions to this. If Brian Ribeiro (2002) is correct, then there are times when I can't tell an epistemically rationalising story about my belief and yet still have the belief in question.

The second thing to note is that the third-person adjudicator would need objective knowledge in order to determine which MPS=Ts are beliefs and which are MPS=Ts.

Let's consider a case where it may seem easy to differentiate between someone's MPS=T and a belief. Imagine someone forms the MPS=T "there is a glass on the table" as a result of an optical illusion. From his/her perspective the MPS=T is a belief. However, we know about the optical illusion. In this case we think we are able to tell the correct etiological story of the MPS=T and can say that that is all it is.

But how do I know that the person's MPS=T has in fact been caused by an illusion? Perhaps, there is a glass on the table behind the illusionary one that is the cause of the MPS=T in question, in which case the MPS=T actually is a belief. Or perhaps the evidence that I have for the claim that the glass on the table is an illusion is in fact the actual trick being played. Can I be sure that you're telling the truth? Can I be sure that you aren't actually deluded? Can I be sure that what I'm hearing you say is actually what you're saying (perhaps I have some sort of auditory disorder that makes me hear things that aren't there)? Perhaps we are brains in a vat, cleverly fed information by a computer that flawlessly mimics reality. At what point does the sceptical inquiry stop? And at what point do I feel satisfied that I know what actually caused your MPS=T? The problem is that we can increasingly raise the level of sceptical doubt about the cause of the MPS=T. The descent into the rabbit hole of metaphysical inquiry, it seems, is unavoidable.

Determining whether an MPS=T was caused in the right way so as to constitute a belief may well be beyond the ability of anyone. I can't make the call with regards to my own beliefs. With regards to making a call about yours, even if I ignore the sceptical concerns, it is impossible for me to know if your very ordinary beliefs have been caused in the right way. You report to believe that there is a glass on the table. While I can see that there is a glass on the table, I can't be sure that this was the actual cause of your belief, or that you actually have a belief.

If this is right, then we end up with the following scenario: we know what the conditions are to satisfy calling an MPS=T a belief, but we can never know if anyone has actually achieved them. If we can never know of either ourselves or of anyone

else whether or not what we or they have is a belief, then it seems we either abandon Davidson's (1986) condition, or stop talking about beliefs. The consequence is either accept that the range of beliefs is so broad as to include delusions, hallucinations, motivated beliefs, *idées fixe* and so on, or stop talking about beliefs at all.

Given these two options, I choose the former. While ruling out things like motivated beliefs and so on may be useful for a truth-belief account, letting these in is preferable to stopping the conversation before we start.[17]

3.2.2. Fantasies, phantasies, delusions and beliefs

Velleman's (2000) objection to my suggested scope of beliefs is that part of what makes a belief a belief is its responsiveness to evidence. As he notes:

> … belief must be characterized by its responsiveness to input as well as its power to produce output … that the input constraints definitive of belief are designed to yield beliefs that are true. (Velleman, 2000:277)

Velleman (2000) is not saying that beliefs must be True to count as beliefs. He acknowledges that we can and do get it wrong (Velleman, 2000:252-255). But, he states that what distinguishes beliefs from other mental propositional attitudes is the fact that they regulate for truth. While I will say more about beliefs regulating for truth below, what's important for this section is that Velleman limits the scope of belief only to those MPS=Ts that regulate for truth. Thus, the person who claims to be Napoleon (even after being faced with counter evidence) is not in the grips of a belief, but instead has a phantasy or a delusion.[18]

---

[17] Of course, the inability to determine whether an MPS=T is a belief is not an objection to Donald Davidson's (1986) metaphysical concern. Davidson may well be right to say that an MPS=T is only a belief if it is caused in the right way, even if we are epistemically constrained in terms of ever knowing which are which. That we can, perhaps, never know if the MPS=T counts as a belief, doesn't undermine Davidson's metaphysical concern.

[18] David Velleman (2000) makes a distinction between fantasies, which are instances of an act of imagination "purposely formed by the subject ... [and] are self-consciously unrealistic, in that they are accompanied by an occurrent awareness of their being in tension with the truth", and phantasies: "wishful imaginings that are not self-consciously unrealistic ..." (Velleman, 2000:266).

Andrew Huddleston (2012) explains:

> Take the case of Joe living in 1990s Cincinnati who purports to believe that he is Napoleon Bonaparte. He talks about his stunning victory at Austerlitz, the brutal Russian winter, his faithful horse Marengo, and his tragic mistakes at Waterloo. Yet this alleged belief about being Napoleon does not seem to aim at the truth, nor to be responsive to fact and evidence. For unlike a simple mistake … Joe's belief flies in the face of all the other facts for which Joe, let us suppose, has ample evidence and which Joe also purports to believe: that he was born in 1950 in Akron, that he works at a sausage factory in Cincinnati, that he has never been to Europe. According to Velleman, such a delusional person does not actually have the belief that he is Napoleon, even if he says that he believes this. It is a fantasy (or ''phantasy'' in Velleman's terminology), not a belief, strictly-speaking. (Huddleston, 2012:214)

The epistemic purist in me would be delighted if Velleman (2000) was right. It would mean that I could simply write off Joe, as well as anyone else whose persistent beliefs I disagree with, as delusional. Thus, the bigot, who like Joe continues to hold onto his/her belief "in the face of all the other facts" (*ibid*.), doesn't have a belief that I ought to get my moral and epistemic knickers in a knot about, he/she is delusional. I can treat the bigot as I would with any other person with a delusion, as needing shock therapy perhaps, and carry on with my day. However, I don't think limiting the scope of beliefs in this way is correct.

Let's first spell out is the difference between a lived fantasy and the type of phantasy/delusion that Huddleston (2012) is talking about.

There is an anecdote of actor Cameron Diaz meeting her co-star Daniel Day-Lewis (a famous method actor) on the set of Martin Scorsese's *Gangs of New York*. Day-Lewis' character in the movie is a violent, one-eyed man called Bill "The Butcher" Cutting.

As part of his preparation for the movies he acts in, Day-Lewis is famous for "being" the character (sometimes for several months) before the movie starts. The story goes that when Diaz went to visit Day-Lewis on set she met instead Bill "The Butcher" Cutting, an experience that she is reported to have found incredibly unsettling.

In this instance I take it that Day-Lewis is in the grips of a lived fantasy that would, given Day-Lewis' reputed devotion to his craft, maintain in the face of counter evidence. A lived fantasy, moreover, that I am happy to say is not a belief. If Joe's delusion that he is Napoleon is nothing more than a lived fantasy, then he too does not have a belief.

However, I take it that the case that Huddleston (2012) is describing (the one that Velleman (2000) wants to rule out as an instance of belief) is not of Joe the method actor, Joe the con artist, or Joe the person with a lot of time on his hands and a vivid imagination. Instead we're talking about Joe who genuinely thinks he is Napoleon (a phantasy or delusion).

In the case of a lived fantasy the fact that the mental state is evidence resistant is not what separates it from belief. What distinguishes the two is that while Day-Lewis may act as though it's True that he is The Butcher, he doesn't think it's True in a way that suggests that he has a belief. In this case, while Day-Lewis is thinking it True that he is The Butcher, he is thinking it True for an extrinsic reason that is transparent to him, thereby ruling it out as a belief. In the case of a phantasy or delusion, however, one of the features of this mental state is that the person thinks it True in a way that suggests a belief.[19]

A clinical psychologist colleague of mine tells me the following story:

> I remember an interesting schizophrenic client I had. He was convinced one of the nurses was stealing his Haloperidol pills in order to keep him psychotic. He told me that he would take note of what was in his pill box early in the morning after night staff had

[19] Not all delusions are going to be things that an agent takes as True. One can imagine a case in which someone acknowledges that what they are seeing and hearing is as a result of a chemical imbalance in his/her brain, rather than something that they take to be True in the sense required to suggest a belief.

> filled it, but when this specific day nurse brought him his pills there were always two yellow pills missing. At one point, his delusion even had me convinced. I checked with his Doctor who said he should be taking two yellow Haloperidols as part of his regimen. She was a mean-spirited nurse too, and his paranoia was getting more dysfunctional by the day, so everything pointed to it being true. I almost opened up a case against the nurse, only to find out he was on generic pills [which were white].

The story above highlights two key points relevant to the discussion. The first is that this mental state looks like a belief. The agent takes the proposition as True and does so for no extrinsic reason that's transparent to him, other than because he takes it as True. The second is that part of what makes this delusion a candidate for belief is that the agent in the grips of a delusion is able to tell an epistemic story about the mental state in question (in this case evidenced by the patient's careful note-taking of his medication). This is a key feature of belief and so if we want to separate delusions from belief we need to look at something other than the explanation that an agent gives for his/her mental state.

Velleman's (2000) criterion to separate delusions from beliefs is that the former are evidence resistant, such that they don't regulate for truth, while beliefs respond to evidence and regulate for truth.

I argue that this criterion cannot be used to separate beliefs and delusions into two separate types of mental state. While I think we can justifiably talk about delusional beliefs, these are dysfunctional instances of the same type of mental state.

Let's consider two beliefs that I have: the computer I'm working on is a Toshiba; and my wife is a living human being rather than a cleverly constructed, computer-controlled automaton.

Speaking for myself, the former is more likely to change in the face of counter evidence than the latter. If someone peels back a sticker revealing a different manufacturer on my computer I'm likely to change my belief in the face of this

evidence. However, should someone peel back my wife's skin to reveal a metal skeleton I would assume either that this is a very clever, if rather twisted, magic trick that my wife is involved in, or that my wife is elsewhere and a clever dummy has been used to replace her. While the former belief would, I suspect, change rather quickly, the latter would take a great deal of convincing. What this shows is that some beliefs are more resistant to counter evidence than others. And, consequently, some beliefs are less likely to regulate for truth than others.

Why some beliefs change quickly while others don't is not a question that needs answering here. What is important here is that some beliefs are more evidence resistant and less likely to change in the face of counter evidence than others. If this is true, then beliefs admit to a degree of evidence resistance on a sliding scale. To use the fact that beliefs regulate for truth as a defining feature of belief must then only be understood in conjunction with this claim.

In addition, what an agent takes as evidence against his/her belief will depend on the agent. So, part of the reason why our beliefs might not change in the face of counter evidence is due to what each agent takes as evidence against his/her belief.

Consider the following example. I take it that the ANC's Twitter feed about the benefits of the e-tolling system in Gauteng is nothing more than propaganda. For me this is obviously true. John, however, disagrees. He thinks that the ANC's Twitter feed contains nothing but facts.

If I present what I take to be overwhelming reasons for my belief to John I expect his belief to change. However, if his belief doesn't change does he have a delusion? No.

To begin with, there is no guarantee that what I take as evidence is what John takes as evidence. If he disregards my evidence as Truth-directed, then his belief has no reason to change.

Similarly, I take it that someone who suffers from a delusion may simply disregard the evidence he is presented with as being legitimate.

In the case of the schizophrenic above my colleague tells me that the patient "warped the counter evidence" when it was presented to him, turning it around in order to tell an epistemically justifying story.

> He was convinced that the nurse was poisoning him. And he had "evidence" in his head of how the nurse was stealing his tablets. I confronted the nurse in front of him: asking her if this was all the medication he needed. The nurse said "yes, I don't choose the medication. I just follow what was in his script". She showed me the script and that corresponded with his medication. When I showed that to him he said "no, she manipulates the file". Everything I presented to him as counter evidence he misinterpreted so that it would fit in with his delusional belief.

So far, both John and the patient above are in the same position. They both take some proposition as True in a way consistent with a belief. Both tell an epistemically rationalising story about it. Neither John nor the patient has any reason to change his mental state from his perspective. Both MPS=Ts are evidence resistant. So what makes one a delusion and the other a belief? The evidence-resistant nature of delusions seems a weak claim to hang this distinction, given that there are MPS=Ts we would want to label beliefs that share this feature.

What further undermines Velleman's (2000) distinction, and suggests that delusions should count as beliefs, is the fact that on occasion delusions and phantasies can regulate for truth. Perhaps the most famous case is of mathematician, John Nash, who battled with schizophrenia and its concomitant delusions.[20] Nash, it seems, at some point was able to recognise that his delusions simply did not fit reality and stopped thinking his delusions represented the Truth (Nasar, 1994:2002). It's not that the delusions disappeared, but rather that he stopped "believing" them. This looks a lot like someone's belief changing in the face of counter evidence.

Sylvia Nasar (1994:2002) writes that while "[t]he actual story of Nash's miraculous triumph over schizophrenia is much more convoluted than the screen version

[20] The story of John Nash's life – in particular his battle with the delusions common to certain types of schizophrenia – was made into a well-known movie, *A Beautiful Mind*.

[suggests] ... the film's depiction of his disease and re-emergence is substantially true" (Nasar, 2002).

While I think we can rightly and accurately say that the schizophrenic patient, above, and Joe have delusional beliefs, there doesn't seem to be anything that stops these from being beliefs. To claim a distinction based on the capacity to respond to counter evidence leads to a slippery slope that is in danger of rendering all our beliefs delusions.

One option for Velleman (2000) would be to say that the counter evidence against Joe and the patient is overwhelming.

Now, if Joe and the patient continue to disregard the evidence against them – continuing to tell an epistemically rationalising story – then who adjudicates a) what counts as evidence and b) when that evidence is overwhelming?

Starting with a). Velleman (2000), it seems, must claim that we have a much better handle on reality than Joe. Because of this we have a much better idea of what should count as evidence for Joe. But what gives Velleman the ability to determine that? If we are all being deceived by an evil demon, for example, then none of our evidence for any of our beliefs counts as evidence.[21]

While it may seem obvious to us that everything that Joe encounters counts as evidence against his MPS=T that he is Napoleon, how do we know how Joe is interpreting the evidence? If Joe is able to tell a convoluted yet epistemically rationalising story about his MPS=T, and if Joe sees this evidence as compatible with the Truth of *p*, and takes *p* as actually being True for no extrinsic reason that's transparent to him, then it seems he has a belief that *p*.

[21] Having been quite clear that my project comes from first-person, experiential perspective, there may be some concern here that I'm now appealing to a metaphysical claim to prove my point. However, given that from Joe's perspective his delusion *just is a* belief, I take it that Velleman (2000) is appealing to something outside of this frame of reference. It seems legitimate, therefore, to highlight that once we do so he opens himself up to this type of sceptical enquiry.

Velleman (2000) may object that in the case of Joe the evidence against his MPS=T is overwhelming. This leads us to b).

There's a concern that we may well be on yet another slippery slope. At what point do we want to say that the evidence against a belief is overwhelming? While the case of Joe seems fairly clear-cut, how do we determine when something goes from a belief to a delusion using Velleman's (2000) distinction? How, for example, would we determine at what point I cross the line from belief into delusion (and thus from belief into an MPS=T) when I claim that my wife is a human being? Is this based on convention, such that if three out of five outside observers think the evidence against me is overwhelming my MPS=T stops being a belief and becomes instead a delusion?

While being labelled as delusional or having a delusional belief may be a matter of convention and have an impact in terms of one's personal freedoms, convention doesn't change the *type* of mental state you have, just the flavour.

Velleman (2000) claims delusions are evidence resistant, while beliefs are not. But this doesn't seem like a good distinction for two reasons. One, MPS=Ts that we would happily call beliefs can be maintained in the face of counter evidence; not all beliefs regulate for truth. Two, delusions can be sensitive to counter evidence. In the case of Joe, he disregards the evidence against him precisely because he doesn't think it's True. But, should Joe, like Nash, see that the evidence against him is true, we are likely to see his delusion regulate for truth. Transferring the adjudicator of truth from the first to the third person may be pragmatically useful in terms of society, but cannot count as a legitimate way of creating a distinction between a belief and a delusion.[22] It seems that if Joe is delusional he is so in virtue of having delusional beliefs.

Being able to explain how these beliefs exist given that "beliefs aim at truth" (Williams, 1973:136) is problematic and one I think other accounts of the truth-belief

[22] There may be a worry here that I'm appealing to a metaphysical claim to back up an experiential one. However, I take it that even from an experiential perspective that while convention may class my mental state as being in line with societal norms or outside them, it couldn't change what my mental state is. My imagining a pink elephant doesn't become a belief just because everyone else says so. It may become unhealthy or silly as a result of convention, but nevertheless it still would remain an instance of imagining while I still had it.

relationship struggle with. An explanation of this feature of belief is discussed in Chapter 5.

3.3. Beliefs can be both conscious and unconscious

I take it that we have both reflective/conscious beliefs (propositions that we are consciously thinking True, occurrent beliefs) and pre-reflective/unreflective/unconscious beliefs (propositions that we are not at that moment consciously thinking True, dispositional beliefs).

I take it that these dispositional beliefs can fall into two types of beliefs: longer-lasting background beliefs and shorter instances of belief that we aren't consciously aware of forming or discarding. An example of the former would be something like knowing where the fridge is in your house, while an example of the latter would be 'there is a bird in the garden'. As Willard Quine and Joseph Ullian (1978) put it:

> [A] belief … may last long or briefly. Some beliefs …. We shall probably retain while we live. Some, like our belief in the dependability of our neighbourhood cobbler, we may abandon tomorrow in the face of adverse evidence. And some, like the belief that a bird chirped within earshot, will simply die of unimportance forthwith. (Quine & Ullian, 1978:10)

In both cases while these beliefs can be at the forefront of one's mind, one can equally have them as unconscious beliefs. Quine and Ullian (1978) note:

> One's repertoire of beliefs changes in nearly every waking moment. The merest chirp of a bird or chug of a passing motor, when recognized as such, adds a belief to our fluctuating store … quickly acquired and as quickly dropped. (Quine & Ullian, 1978:9)

Technically, the quote above is about the fact that many of our beliefs are temporary. However, it does suggest that many of our beliefs are formed and revised

unconsciously. I take it that if we constantly had to consciously make sense of all of our beliefs we would stop being able to do anything else.

Further, Shah and Velleman (2005) discussing cases of doxastic deliberation note that "the norm of truth that controls doxastic deliberation needn't control other [belief-forming] processes" (Shah & Velleman, 2005:501). This suggests that in addition to cases of doxastic deliberation, where the process of belief formation is done in full consciousness, there are beliefs that are formed, revised and maintained unconsciously. If there are beliefs that are formed and revised unconsciously, then it is no great leap to think that we have beliefs that aren't in the forefront of our minds.

This feature of belief is something I'm going to more-or-less stipulate with three clarifying remarks.

The first is that I'm not making some sort of ontological, behavioural-dispositional claim about beliefs such that "for someone to believe some proposition *P* is for that person to possess one or more particular behavioural dispositions pertaining to *P*" (Schwitzgebel, 2010). I'm not claiming that in order for S to have a belief that *p* it is a necessary condition that S must behave in certain ways "pertaining to *P*" (Schwitzgebel, 2010). Rather the claim is that in cases of turning on taps, televisions etc., these behaviours can be driven by dispositional beliefs.[23] The behaviour, therefore, is merely indicative rather than constitutive of a belief.

The second is that I am not talking about belief attribution either. Making belief ascriptions about others is notoriously difficult. So when I see someone behaving in a certain manner I might think that he/she is doing so as a result of a certain belief, interacting in a certain way with a specific desire and the agent's will so that he/she performs the action. In these cases while I may be correct in labelling the belief from which the agent is acting I can never be certain. The actor, Day-Lewis, may act as

[23] For an overview of dispositional accounts of believing as well as an overview of the ontological theories of belief see Eric Schwitzgebel (2010).

though he believes he is The Butcher, but not actually have these beliefs. In this case, the motivation for his action comes from his lived fantasy.[24]

The third is that with regards to attributing a belief that *p* to oneself I take it that one has to, at the time of making the ascription, be consciously thinking "that *p* is true"/"that *p*"/"*p*", or something sufficiently similar to satisfy the conditions necessary to fulfil the belief attribution.

However, the claim here is not about attributing beliefs to oneself or another. The claim here is that we can make sense of an agent having a belief that *p* in spite of the fact that the agent is not consciously thinking "it is true that *p*"/"*p*", or any such variation.

That we can have both occurrent and dispositional beliefs, as well as the fact that we can form and revise beliefs both consciously and unconsciously, is important for two reasons. The first, as I argue below, is that I maintain that features of belief like transparency – which Shah (2003) and Shah and Velleman (2005) think apply only to deliberative beliefs – apply to beliefs, *tout court*. An account of the truth-belief relationship, therefore, needs to be able to account for the features of belief as they apply to both conscious and unconscious beliefs. The second is that one of the things that makes the fact that beliefs regulate for truth so interesting is that they appear to do so outside of our will. In other words, we don't consciously choose for our beliefs to regulate, they just do. Further, in regulating for truth, our beliefs, conscious or unconscious, appear to follow epistemic norms. Explaining this is one of the key features an account of the truth-belief relationship has to explain.

3.4. Epistemic reasons are the only reasons that count as reasons for belief

Now, irrespective of whether your belief is conscious or unconscious the only reasons you can see yourself as having for a belief are epistemic reasons.

[24] It is important to note that I'm not taking a stand on the belief-desire account of action. Velleman (2000) provides a nice account of behaviour that stems from imagining or make believe (Velleman, 2000:256-263). I like Velleman's account and think it has merit. However, it is not germane to my project to defend or refute his position. My point is simply to point out that behaviour is indicative of, but not a requirement of, a belief. As a result, we can use behaviour as evidence of another's belief, though it is by no means a certainty that the behaviour is reflective of a belief.

This isn't a moral point that agents should believe in line with Truth (though we might want to argue that this is also true), but rather when you're weighing up reasons to believe that *p* you can only take epistemic reasons as reasons to believe that *p*.

This feature is entailed by what a belief is. Because to believe *p* is to take *p* as True any reasons that you judge fail to provide evidence of the proposition's Truth cannot be a reason to believe that proposition. Put more simply, any reason that is not an epistemic reason for *p* does not and cannot count as a reason to believe that *p*.

As Thomas Kelly (2003) explains, trying to believe for non-epistemic reasons just doesn't work:

> Suppose, for example, that I can save my life by holding some epistemically irrational belief. Suppose further that the belief concerns some subject matter with respect to which my having true rather than false beliefs is a matter of relative indifference. Now, if epistemic rationality just is instrumental rationality, then I think we can safely conclude: all things considered, it is rational for me to hold this belief. (Kelly, 2003:620)

But, as Kelly (2003) concludes, we can't form our beliefs as a result of pragmatic reasons, no matter how prudentially useful they may be.[25]

There is an important caveat here. It's not that we can only take epistemic reasons to believe that *p*, but rather that we can only take *what we take* as an epistemic reason as a reason to believe that *p*. If, for whatever reason, an agent takes a pragmatic reason as a convincing epistemic reason for the Truth of *p*, then he/she will form the belief that *p*.[26]

[25] While Thomas Kelly's (2003) aim (to show that epistemic reasons are not merely a subset of instrumental reasons) is different to the point I'm making here, these examples do tease out this doxastic constraint.

[26] For the sake of convenience I use the phrase "epistemic reason" and "non-epistemic reason" respectively to mean "what one takes as an epistemic reason" and "what one takes as a non-epistemic reason" from here on unless spelling out the distinction is important.

Thus, the claim here "is not that there are no cases in which a desire or some other psychological state … might produce the belief that p" (Scott-Kakures, 1994:78), as Dion Scott-Kakures (1994) puts it, but rather that we can't choose to believe something for what we see as instrumental reasons, no matter how compelling.

Richard Feldman (2000) notes:

> Suppose that one belief is prudent for me – it will maximize my well being or what I care about or what is important to me – but it is not a belief I epistemically ought to have since I lack evidence for it ... I can see no values to which we could be appealing when we ask whether the overall prudential benefit trumps the epistemic cost ... There is no meaningful question about whether epistemic oughts "trump" or are trumped by other oughts. (Feldman, 2000:693-694)

If epistemic reasons were like pragmatic reasons we should be able to abandon an epistemic reason in favour of a pragmatic reason. But this isn't possible. Epistemic reasons are a different type of reason, and the only reasons we can take as reasons for a belief are epistemic reasons.

Of course, there may be ways of deceiving yourself into taking some non-epistemic reason as an epistemic reason. I can, for example, try to convince myself that the evidence I'm confronted with of my son's stupidity is actually a sign of a special type of artistic genius. While this may not be initially successful, I can start to read less-than-reliable websites and journals that support my claim, eventually convincing myself that these purported reasons are in fact good epistemic reasons. However, this type of project of self-deception supports rather than detracts from the notion that the only reasons for belief are epistemic reasons. In trying to deceive myself I am, on some level, acknowledging that I need to have epistemic reasons in order to form a belief.

This appears to be true of both our conscious and unconscious beliefs. When looking for your keys, for example, you're probably not consciously thinking "I don't know where my keys are". Instead you are engaged in the project of trying to find them and

your belief has faded into the background while you look under couch cushions and search your pockets for the fifth time. Your belief that you don't know where your keys are, however, doesn't change *because* you want to find them. Should you catch sight of them your belief is likely to change in the face of this epistemic reason. This, mostly, isn't a conscious case of going "Seeing my keys provides me with an epistemic reason to believe that I know where my keys are. I find this reason epistemically compelling. I'm now going to believe I know where my keys are". In this case your belief that your keys are lost disappears automatically and unconsciously; and disappears only because you've got what you take as epistemic reasons for believing that your keys are found. Your desire to find them, no matter how pressing, simply cannot be something you take as a reason to believe.

This feature is important for two reasons. The first is that any adequate account of the truth-belief relationship needs to be able to explain this feature. The second is that, as we shall see later on, not only are epistemic reasons the only reasons we can take as reasons to believe, these reasons somehow motivate our beliefs to change in line with truth.[27]

3.4.1. Epistemic reasons, epistemic explanations and second-order beliefs

There are three closely linked corollaries to this feature of belief. The first is that to have a belief that *p* means telling an epistemically rationalising story about *p*. The second is seeing that your belief that *p* is based on or sustained by non-epistemic

[27] Carl Ginet (2001) argues that we can decide to believe that *p* for prudential reasons in some special instances where a) the evidence for the truth of *p* is at least as strong as the evidence for its falsity, and b) believing that *p* would be prudentially beneficial. One of Ginet's examples is the following:

> Before Sam left for his office this morning, Sue asked him to bring from his office, when he comes back, a particular book that she needs to use in preparing for her lecture the next day. Later Sue wonders whether Sam will remember to bring the book. She recalls that he has sometimes, though not often, forgotten such things, but given the inconvenience of getting in touch with him and interrupting his work and the thought that her continuing to wonder whether he'll remember it will make her anxious all day, she decides to strop fretting and believe that he will remember to bring it. (Ginet, 2001:64)

While there is a lot more that needs to be said to undermine Ginet's (2001) argument the short version of my response is as follows: I believe that Sue can decide not to fret about whether Sam will bring home the book, but I don't believe that this is because she has come to believe that Sam will bring home the book. Stopping to fret does not entail having a certain belief about something. All it means is that one has stopped worrying about it. Ginet's argument, therefore, to my mind doesn't challenge the claim that only epistemic reasons count as reasons to believe.

reasons means giving up your belief that *p*. And finally, the third is that you cannot believe that *p* and believe that your belief that *p* is sustained or explained non-epistemically. Barbara Winters (1979) defends the latter two, while Ward Jones (2002), more explicitly, defends all three.

Let's examine these in order.

One, when you examine your belief that *p*, if you are being honest with yourself, you will find that in explaining why you believe that *p*, the reasons you give will be epistemically rationalising. What's interesting is that this constraint is so strong that in some cases the agent has to stretch the bounds of credulity to meet it. As we saw with the schizophrenic patient above the more evidence that went against his belief the weirder and more wonderful his epistemically rationalising story became, but it was still an epistemically rationalising story. The patient didn't say, "I know I have no epistemic reasons to believe this, but I still believe that my nurse is trying to harm me". Instead, he was forced, as it were, because he still had the belief, to tell an epistemically rationalising story about his belief. In this case, about how the counter evidence that was presented against his belief wasn't *really* counter evidence.

Two, if you find that you can't or don't have an epistemic reason for your belief, you give it up. The claim here is not that we would continue to hold a belief, but feel obligated to give up, but rather that the belief would cease to exist upon us recognising that it was non-epistemically constituted. As with the first claim this follows naturally from the concept of belief. To see that you have no epistemic reasons for your belief that *p* means you have no reason to think that *p* is True. As a result, given the direct link between belief and truth, your belief that *p* ceases to exist in the face of this acknowledgement.

The third claim is that one cannot believe that *p* and believe that one's belief that *p* is sustained or explained non-epistemically. The claim here is not that we cannot have non-epistemically caused beliefs. Nor even that I can't say that I have non-epistemically caused beliefs generally speaking. Simply, the claim is that when I attend to *a* belief I cannot *believe* of that belief that it is non-epistemically explained. As with the other two claims, this claim seems to follow from the concept of belief. In

fact, Winters (1979) and Jones (2002) argue that given the concept of belief we are conceptually constrained from believing that *p* and believing that our belief that *p* is based on non-epistemic reasons. To do so, they argue, implies a logical contradiction.[28]

I have previously argued (Viedge, 2004) that Jones (2002) and Winters (1979) fail to prove that the constraint on our beliefs is a conceptual constraint. I argued that in the admittedly exceptional case of finding an argument both unanswerable yet unbelievable we do sometimes have second-order beliefs that appear to contravene this constraint.[29] However, despite my argument that the constraints shouldn't be seen as conceptual, I do think that Winters (1979) and Jones (2002) describe a ubiquitous, *ceteris paribus*, limitation on what we are able to say of and believe about our beliefs. When I examine a particular belief and consider why I believe it, I cannot, in all but the most unusual of cases, have certain second-order beliefs about that belief. In addition, I can't have a belief that I take to be based on or sustained by non-epistemic reasons. Finally, I can't look at one of my beliefs, provide a non-epistemically rationalising story about that belief and still believe it.

3.5. Transparency

Shah (2003) notes that when one is deliberating about "*whether to believe that p* … this question is settled by, and only by, answering the question *whether p is true*" (Shah, 2003:447). The former question is transparent to the latter one.

---

[28] The argument, in brief, is as follows. To believe that *p* is to take *p* as True. To believe that your belief that *p* is sustained or formed non-epistemically is to essentially claim of *p* that it is not true. Thus, to believe of your belief that *p* that it is based on or sustained by non-epistemic reasons is to believe that *p* and believe that your belief that *p* is false. This is a logical contradiction and is thus conceptually impossible.

[29] How this is cashed out is explained in exhausting detail in my MA dissertation. The very short version is that finding an argument unanswerable yet unbelievable would be to accept that a certain argument is sound and yet find oneself unable to believe the conclusion. Given that we cannot get round the argument, so to speak, and yet still have a belief in the face of this argument we must conclude that our belief is based on some non-epistemic reason. David Hume (1977) is perhaps the first to bring this notion to light in *An Enquiry Concerning Human Understanding*, arguing that certain sceptical arguments cannot be believed, despite the fact that the arguments cannot be answered. Most famously he argues that induction is a flawed system of reasoning and then seemingly finds himself in a position where he himself cannot believe the conclusion of his own argument (Hume, 1977:15-20). This notion, as well as some of its consequences, is nicely spelt out in Ribeiro's "Epistemological Skepticism(s) and Rational Self-Control".

What this means is that when considering 'whether to believe that *p*', we acknowledge that this question is solved only by providing reasons that answer 'whether *p* is true', or simply 'whether *p*'.

I think there are three ways of understanding transparency.

The first interpretation of what transparency could mean is that transparency is simply the acknowledgement that the link between belief and truth is direct. If 'whether to believe that *p*' is transparent to 'whether *p*', this appears to acknowledge that the link between belief and truth is direct. If you take *p* as True, you believe that *p*.

What backs up this interpretation of transparency is the quote I used above by Shah (2003) to discuss the direct link between belief and truth.

> Truth is not an optional end for first-personal doxastic deliberation, providing an *instrumental* or *extrinsic* reason that an agent may take or leave at will. Otherwise there would be an inferential step between discovering the truth with respect to *p* and determining *whether to believe that p* … But there is no such gap. (Shah, 2003:447)

If this is what Shah (2003) means by transparency, then I am happy to accept it as it simply confirms what I already take to be a feature of belief and furthermore is a feature of both our occurrent and dispositional beliefs.

Alternatively, transparency is a claim about the nature of the reasons that can count as reasons for belief. If transparency is a claim about the nature of the reasons that can count as reasons for belief, then the second way of understanding transparency is that in deliberation only epistemic reasons can count as reasons for belief. If this is what transparency means, then deliberation simply highlights that epistemic reasons are the only reasons that count as reasons for a belief. This certainly seems like a fair characterisation of what Shah (2003) means by transparency. Shah's contention is that 'whether to believe that *p*', is solved only by providing reasons that answer 'whether *p*' and this is a clear example of how only epistemic reasons count as reasons for

belief. As with the first way of understanding transparency this is something I already accept as a feature of both occurrent and dispositional beliefs.

The third way of understanding transparency is that in instances of doxastic deliberation it is not just that the link between belief and truth is direct, but that when deliberating there is a direct link between epistemic reasons and the belief under consideration. So it's not just that epistemic reasons are the only reasons for belief. It's that when I encounter an epistemic reason for or against *p* I don't need to weigh this reason against instrumental reasons because the epistemic reason speaks directly to what I'm deliberating about in a way that prudential reasons shouldn't. If I find the reason compelling I form the belief, if I don't I reject it.

As with the two previous interpretations, I'm happy accepting this interpretation of transparency. If transparency is to be understood as a claim that epistemic reasons have a direct link with belief such that there is no intermediary gap between seeing *e* as an epistemic reason and taking *e* on board in terms of how it relates to whether to believe that *p*, then this is equally acceptable. This seems to be a logical consequence of the direct link between beliefs and truth and the fact that epistemic reasons are the only reasons that count as reasons when it comes to beliefs.

Now, while I'm content to accept transparency as a feature of belief if it is understood in any of the above three ways, there is one important caveat. As with the primacy of epistemic reasons, above, it must be what an agent *takes* as an epistemic reason that counts rather than what actually is an epistemic reason. If Shah (2003) and Shah and Velleman (2005) think that in deliberation only actual epistemic reasons count towards whether to believe that *p*, then this would require deliberation to be never be swayed by non-epistemic reasons. Since this is obviously not the case, it must be the case that it is only what the agent *takes* as an epistemic reason for *p* that counts towards whether to believe that *p*.[30] Similarly, if transparency is a claim about the direct relationship between belief and truth, then it must be what an agent takes to be True, rather than what is actually True.

[30] I argue for the claim that non-epistemic factors can shape even deliberative beliefs obliquely below. However, it seems obvious that someone with a predisposition to dislike a certain group, for example, is less likely to take testimony from someone from that group as veridical. In this instance, we have a non-epistemic factor shaping the way that someone sees evidence.

Shah (2003) is aware of this concern. As he notes:

> This point about transparency does not imply that non-evidential factors cannot causally influence belief, or that we cannot recognize that such factors may infect our thinking, but rather it implies that we cannot recognize the relevance of such factors in our deliberative reasoning about what to believe. (Shah, 2003:463)

With this caveat in mind, transparency seems like an obvious feature of belief to accept. It either cashes out the features already detailed above (in 3.1 or 3.4), or it details a feature that follows logically from these two features.

Shah (2003) and Shah and Velleman (2005) make two claims about transparency.

The first is that not being able to account for transparency counts against teleological understandings of the truth-belief relationship and that because a teleological account cannot account for transparency we should abandon it in favour of a Normative Account. The inference, therefore, is that transparency is feature of belief and, therefore, ought to be explained by the truth-belief relationship.

The second is that transparency is only a feature of doxastic deliberation. Shah (2003) and Shah and Velleman (2005) argue that while a teleological account cannot explain transparency – which is evident in instances of doxastic deliberation – a teleological account "seems to explain why evidence-driven processes such as perception normally induce belief" (Shah, 2003:473). The claim is that while we need a normative account to explain doxastic deliberation, we can keep a teleological account for other belief-forming processes.

In what follows, I argue that one of these two claims about transparency must be wrong: either transparency is a feature of deliberation or transparency is a feature of beliefs, *simpliciter*. If it is a feature of only deliberation, then it is not a feature of belief. If it is not a feature of belief, then an account of the truth-belief relationship does not need to explain it. Thus, if a normative account explains only deliberation it

is not an account of the truth-belief relationship. While an account of deliberation must be able to explain transparency, an account of the truth-belief relationship need not. If, however, transparency is a feature of belief, then it is a feature of belief, *simpliciter*. If transparency is a feature of belief, *simpliciter*, then an account of the truth-belief relationship must be able to explain it *as well as* the other features of belief we have looked at. For if a theory explains only one feature of belief, leaving the other features to be explained by another account of the truth-belief relationship, it is at best an *ad hoc* theory of the truth-belief relationship.

I argue that while Shah (2003) and Shah and Velleman (2005) are right to call transparency a feature of belief, they are wrong to think that it applies only to doxastic deliberation. I argue that in much the same way that Pascal's Wager highlights the primacy of epistemic reasons, doxastic deliberation highlights transparency.

3.5.1. Why transparency applies to more than deliberative beliefs

Let us consider what it means to deliberate.

We can deliberate about all sorts of things. We can deliberate about whether something is good for us from a moral or health perspective. We can deliberate about whether someone is guilty or innocent. And we can deliberate about whether to believe that *p*. To deliberate is to consciously consider some proposition with a specific project in mind.

To deliberate about whether this painting is a good painting, for example, I am consciously considering what I think of the painting before making a judgement about that painting; declaring it a good or bad piece of art. Doxastic deliberation is a conscious project of determining whether to believe that *p*.

This is different from other belief-formation in much the same way that day-dreaming is different to dreaming. In the former, the task is done consciously and with purpose in order to shape the outcome. Whereas in dreaming you're generally a passive recipient of whatever your subconscious throws up at you. Similarly, in cases of doxastic deliberation you're actively deciding whether to believe that *p*, which means

actively deciding whether *p* is True. In other cases of believing you sort of find yourself with beliefs.

If transparency is a feature only of deliberative beliefs it is so in virtue of the intentional stance that the agent adopts towards the proposition.[31] In other words, if transparency is only a feature of deliberative belief it is because of the intention by the agent to determine whether to believe that *p*. If it is not this intention to determine whether to believe that *p* that accounts for transparency only applying to doxastic deliberation, then transparency would be a feature of all beliefs.

But if transparency is as a result of the intentional stance of the agent, then it is not a feature of any type of belief. It is a feature of the project that the agent has that in this instance relates to a belief. The argument against transparency counting only as a feature of deliberative beliefs is as follows:

1. There is a distinction between the features of a project that a believer intentionally engages in with regards to a belief and a feature of belief, *simpliciter*.[32]
2. The intentional stance an agent takes towards a proposition, in and of itself, tells us nothing about the features of belief.
3. If transparency is a feature of doxastic deliberation and only doxastic deliberation, then it is a feature of the project that a believer has engaged in with regards to belief, rather than a feature of belief.

[31] I am not using the phrase "intentional stance" in the same way that that Daniel Dennett (1987) does in his book *The Intentional Stance.* Dennett's intentional stance is the explanatory position we can take towards things such that we ascribe intentions to things. In doing so we imply that there is some mental state that needs to be explained. All I mean here is the deliberate decision to try to believe something – in the case of Pascal's Wager to believe that God exists, in the case of doxastic deliberation to believe that *p* if and only *p* is true. The mental state, the intention, is assumed.

[32] Edmonds (2007), in determining whether beliefs can be said to aim at truth, makes a similar distinction, examining whether the metaphor relates to the aims of an agent or the goal of belief (Edmonds, 2007:15-22). She examines whether we can say of agents that they aim at truth, before examining whether beliefs can be said to have the goal of truth. She dismisses that agents *qua* agents can be said to innately aim at truth and denies that beliefs can be said to have a goal of truth. Her argument is that agents don't seem to be particularly interested in getting to the truth, while the fact that beliefs admit to non-epistemic reasons undermines the claim that beliefs, themselves, have the goal of truth. My concern is different. My argument is that the aims of an agent (even when that agent aims at truth, as is the case in doxastic deliberation) tell us nothing directly about the relationship between beliefs and truth.

4. If transparency is not a feature of belief, then an account of the truth-belief relationship needn't worry about being able to explain transparency.
5. However, transparency is a feature of belief.

Therefore,

6. Transparency is a feature of all beliefs rather than only of doxastic deliberation and an account of the truth-belief relationship would need to explain this and the other features of belief.

The detailed version goes like this. As we have seen, there are certain constraints on how I can view my belief. I can't, for example, see my belief as based on or sustained by non-epistemic reasons in all but the most unusual of circumstances. I find I believe in line with truth, even if what I would prefer to believe is at odds with this and so on. These constraints on what and how we can believe are brought about by the relationship that truth has with belief and are features of belief.

That said, while the truth-belief relationship admits to certain features, which in turn impact on our doxastic lives, this doesn't mean that as a believer I can't engage in doxastic projects that will be at odds with the constraints imposed by truth-belief relationship. In what follows, I examine Pascal's Wager as an example in which the aims of the agent are at odds with the constraints of a belief.

Pascal's Wager is: given the benefits of a belief in God, irrespective of the Truth of such a being's existence, we should try to believe God exists. Blaise Pascal (2005) claims that if religious views of God are correct, then believing in God provides eternal life in heaven. A failure to believe in God on the other hand condemns one to an eternity in hell. Given this we have very good prudential reasons for believing in God, even if it is unlikely that God exists, Pascal thinks we should engage in this project. For if there is the slightest chance that God exists, and believing in God is a prerequisite for avoiding an eternity in hell, then we should try to form the belief that God exists. In this case, Pascal is suggesting that one's belief-forming process, rather than being 'naturally' guided by the evidence, aims at a specific conclusion: namely a belief in God.

Pascal's Wager is a clear case where an agent adopts an intentional attitude towards a proposition, in this case trying to form a belief in God. As part of this project there are certain norms that go along with the project: that one engages in practices that provide credibility to the claims under consideration, for example; or that one adopts a less-than-rigorous epistemic attitude towards certain claims; or that one tries to convince oneself that some non-epistemic reason is an epistemic reason.

These norms developed out of the project of the believer are at odds with the norms of belief. As argued for above, the only reasons we can take on board as reasons for believing something are epistemic reasons. But Pascal's Wager appears to require us to do completely the opposite. Thus, the aim of the believer and the aim of belief aren't always the same thing – even when it comes to the same candidate.

In the case of Pascal's Wager, what does the aim of the believer (to form a belief that God exists) tell us about the aim of belief or the features of belief? As it stands, nothing. That an agent can take a certain attitude towards a proposition, tells us only that an agent can take a certain attitude towards a proposition; in this case, an attitude that appears to contain an element of self-deception. Similarly, the norms associated with this project, *prima facie*, tell us only what we ought to, need to, or must do in order to achieve the intended belief. They are not the norms of belief.

Certainly, Pascal's Wager is not a feature of belief. Nor is the norm that we ought to adopt a less-than-rigorous attitude towards evidence a feature of belief. The former is a project that an agent can engage in that pertains to a belief. The latter is a feature of that project. In fact, none of the norms generated by Pascal's Wager are features of belief. Or, if they are, they are so only coincidentally.[33]

[33] Consider the following. Eating is part of a teleological process. The purpose of this process is to provide energy for the body to function. However, in terms of eating we can eat for all sorts of reasons. We can eat for pleasure, to taste something exotic, to curb depression, to be polite and so on. Now, these aims may tell us some quite interesting things. For example, each of these aims may generate concomitant norms, required to help satisfy the agent's intentional aim. And these norms may be of genuine interest in the examination of the aims in question. But none of these pragmatic reasons tell us anything interesting about the relationship between eating and its inherent purpose; these processes/norms are independent of the agent's intentions. This is true even when an agent eats with the intention of providing energy to the body. In this case, when the agent eats he/she does so in accordance with the inherent relationship between food and the characteristic purpose of eating. However, even if we eat with the intention of providing energy for one's body this tells us nothing new or interesting about the relationship that isn't subsumed by the inherent purpose of eating.

Much like Pascal's Wager, doxastic deliberation is a project that an agent can engage in. If transparency is a feature or norm generated out of this project, then that norm pertains to deliberation, rather than belief; in much the same way that a need for self-deception is a feature of Pascal's Wager, not of belief. If transparency is a feature of deliberation, then the truth-belief relationship doesn't need to account for it. Transparency is a feature generated out of the aim of the agent's project and is therefore not a feature of belief.

However, while transparency understood in this way cannot be a feature of belief, this is not to say transparency isn't a feature of belief. My argument so far is that if transparency is only a feature of doxastic deliberation, then it is not a feature of belief. However, I think that transparency is a feature of belief and therefore applies to all beliefs, not just deliberative beliefs. I argue that doxastic deliberation reveals transparency as a feature of belief in much the same way that Pascal's Wager reveals the fact that we must see ourselves as having epistemic reasons for our beliefs.

As discussed earlier, when an agent forms a belief he/she must do so only on what he/she takes as being epistemic reasons. At first blush this rules out Pascal's Wager as a tenable doxastic project.

However, there may be ways to "dull your mind and make you believe" (Bennett, 1990:88). If one focuses only on the evidence for God's existence, immerses oneself in the lifestyle and surrounds oneself with people who uncritically believe in God you may find that "you [come to] believe quite naturally" (Pascal quoted in Jones, 2002:229).

Now, while Pascal's Wager, in and of itself, tells us nothing about the truth-belief relationship, the processes involved to meet Pascal's Wager reveal that we have to think that the reasons for *p* are epistemic reasons for *p* if we're going to believe that *p*. Thus, the reason we have to go out of our way to try to meet Pascal's Wager is that we have to ensure that we adhere to the conditions that are set by the relationship between truth and belief. In other words, the reason why Pascal's Wager requires a level of self-deception is that we cannot, as a result of the relationship between belief

and truth, see that reasons we have for *p* are non-epistemic reasons for *p* if we are to believe that *p*.

However, while the self-deception is necessary in fulfilling Pascal's Wager because of the pre-existing truth-belief relationship, this self-deception is only a feature of projects like Pascal's Wager rather than a feature of belief in and of itself. What is a feature of belief is that only epistemic reasons for *p* count as reasons to believe that *p*, which is why we need to mislead ourselves if our intention is to form a specific belief.

Turning our attention to doxastic deliberation, my claim is that in much the same way that Pascal's Wager reveals the primacy of epistemic reasons, doxastic deliberation reveals transparency.

Given the argument above, there are two options in terms of how to understand transparency. Either transparency is a norm that is born out of the project of doxastic deliberation (in the same way that self-deception is a norm generated out of Pascal's Wager), or doxastic deliberation reveals transparency as a feature of belief (in much the same way that the primacy of epistemic reasons is revealed in Pascal's Wager).

If transparency is a feature only of deliberative beliefs as Shah (2003) and Shah and Velleman (2005) claim, then they must concede that transparency is not a feature of belief, but instead a feature of deliberation. And if it's a feature of deliberation, then it is not a feature that the truth-belief relationship needs to account for as it is not a feature of belief. Shah and Shah and Velleman's normative account, therefore, is not an account of the truth-belief relationship, but an account of doxastic deliberation.

If deliberation reveals transparency in the same way that Pascal's Wager reveals the primacy of epistemic reasons for belief, then deliberation highlights an already existing feature of belief. If this is the case, then transparency pertains to beliefs, *simpliciter*, rather than just deliberative beliefs.

It seems to me that in whichever way we understand transparency this is a feature of belief, *simpliciter*, which is highlighted when we engage in doxastic deliberation.

Doxastic deliberation is the considering, in full conscious, of a candidate for belief. Unlike Pascal's Wager, where I'm trying to get to a specific belief, I am in cases of doxastic deliberation trying to believe correctly. I am, in a sense, trying to believe in line with what a belief is. It seems that given this project I will take extra caution in trying to consciously meet the epistemic norms a belief appears to inherently follow. In the case of deliberation, the goals of the agent and apparent "goal" of belief are matched onto one another. But this doesn't mean that transparency, however it is understood, is a feature of deliberation. Rather it suggests that deliberation reveals this feature of our beliefs by being a process of belief formation done in full consciousness.

Thus, transparency, for me, should be understood as a feature of belief and is, therefore, something that an account of the truth-belief relationship needs to explain.

3.6. Beliefs regulate for truth

As Shah and Velleman (2005) note beliefs are "regulated for truth" (Shah & Velleman, 2005:498). When an agent sees that his/her belief is not true, the agent ceases to hold that belief. Again, as with the other features of belief, this regulation for truth is a feature of both our occurrent and dispositional beliefs as the quote above by Quine and Ullian (1978) about the ever-changing nature of our beliefs suggests.[34]

I take this feature as uncontroversial for two reasons. One, it seems obviously true that our beliefs regulate in this way. Two, what makes things like *idées fixe* so controversial in that they fly in the face of this norm.

That beliefs regulate for truth highlights not only what Owens (20030 calls "the authority of epistemic norms" (Owens, 2003:284) (what I call the primacy of epistemic reasons), but the fact that these epistemic reasons have such motivational force that our beliefs regulate for truth even unconsciously, or against our desires. An

[34] I use truth, rather than Truth here because beliefs that regulate don't always do so in line with Truth. For example, I might, for some reason, change my belief that the Earth is round in favour of a belief that the Earth is flat. In this case the belief has regulated for truth, but not for Truth.

account of the truth-belief relationship, therefore, needs to explain "not only the authority of epistemic norms but also their motivational efficacy" (*ibid*.).

3.7. Beliefs are involuntary

The fact that beliefs regulate for truth, as well as the direct link between truth and belief, has led some authors, notably Williams (1973) and Scott-Kakures (1994), to think that voluntary belief-formation is conceptually impossible.

Now, as Vahid (2006) remarks, with a beautiful penchant for understatement, "the issue of doxastic voluntarism, as we all know, is quite tricky" (Vahid, 2006:316).

The reason why I engage with this can of epistemic worms at all is that if beliefs are something that we can choose to have at will, then we needn't worry about *idées fixe* or motivated beliefs at all: in these cases the agent has simply chosen to have or keep certain beliefs. An account of the truth-belief relationship, therefore, needn't worry about explaining these kinds of belief. (That said, the account of the belief-truth relationship would have to explain how it is that beliefs are voluntary.)

I argue that while it may well not be a conceptual impossibility, believing at will is a psychological impossibility.[35] This is not because I think there is a particularly compelling argument against the conceptual argument, but because my argument needn't rest on a defence of the conceptual argument.

The argument for the involuntarism of beliefs is generated out two claims. The first is that the concept of belief seems to preclude the ability choose to believe/believe at will (this is the basis for the conceptual argument). The second is that our experience as doxastic agents supports the claim that we cannot believe at will.

[35] Jonathan Bennett argues that it is possible to imagine "a community each of whose members can be immediately induced to acquire beliefs" (Bennett, 1990:93). If Bennett is correct, then the conceptual argument against doxastic voluntarism fails, in virtue of the fact that there is an existing case which disproves this. However, this wouldn't undermine the claim that from our perspective as agents it is still, psychologically, impossible for us to believe at will. I personally believe that Bennett's thought experiment begs the question, assuming that it is conceptually possible to believe at will in order to show that the conceptual argument against believing at will fails. However, as mentioned, I do not think that defending the conceptual argument here is necessary. All I need to show is that believing at will is, from our perspective, psychologically impossible.

Let's consider these in order.

To believe that *p* is to take *p* as being True, not for any extrinsic reason that is transparent to us, but simply because one takes *p* as True. Given this, it seems inconceivable that you could choose to believe something. To do that would be to choose for something to be true and this seems impossible to do.

Consider, for example, being told to believe that the moon is made out of cheese. It is impossible. You could imagine it easily, but it's hard to conceive that you could genuinely believe it. To believe it would be to genuinely take it as True that moon is made out of cheese simply because you want the moon to be made out of cheese. This, to me, is impossible and is so because of what a belief is. Beliefs, therefore, appear to be beyond our voluntary control.

Williams' (1973) argument against voluntarism is straightforward:

> If I could acquire a belief at will, I could acquire it whether it was true or not; moreover I would know that I could acquire it whether it was true or not. If in full consciousness I could will to acquire a 'belief' irrespective of its truth, it is unclear that before the event I could seriously think of it as a belief, i.e. as something purporting to represent reality. (Williams, 1973:148)

The claim here, as Scott-Kakures (1994) points out, is not that beliefs are always formed epistemically, but rather that one cannot choose to form a belief:

> … what is being claimed is not that there are no cases in which a desire or some other psychological state, functioning in some roguish manner, might produce the belief that p. What is denied is something very different, and amounts in essentials to this: I want to believe that p and so will or intend to believe that p, and as a result of my willing to believe that p, I come, in a way appropriate to intentional action, to believe that p. The emphasis here is upon the typical and appropriate connection which we take to hold between willings and the products

> of the will in virtue of which the products of the will can be said to have been produced intentionally. In brief, what we deny when we deny that an agent can believe at will is, as Brian O'Shaughnessy puts it, that "there [can] be an act that is the bringing about of a belief as there is a basic act that is the bringing about of an arm rise". (Scott-Kakures, 1994:78)

The argument has a great deal of intuitive appeal. Given this concept, it is impossible that one could choose to believe something that was false. Further, our experience corresponds with this claim. When I try to believe something at will I find I can't do it. This impossibility is not only a seeming logical consequence of the concept of belief; it is something that resonates with our experience of believing.

This inability to choose our beliefs transcends our practical desires and motivations. As Scott-Kakures (1994) points out, even when we have good practical or prudential reasons to believe something that isn't true we find it impossible.

> George Orwell's*1984* provides us with a useful example of the captivity of the will with respect to belief. Winston, our sort-of-hero, and O'Brien, our sort-of-villain, have a chilling exchange.
>
> > O'Brien held up his left hand, its back toward Winston with the thumb hidden and the four fingers extended. 'How many fingers am I holding up, Winston?' 'Four.' 'And if the Party says there are not four but five – then how many?' 'Four.' The word ended in a gasp of pain.
>
> This is repeated many times. Winston remarks, after O'Brien notes that he is a slow learner, "How can I help it." "How can I help seeing what is in front of my eyes? Two and two are four." (Scott-Kakures, 1994:79)

Winston has excellent practical reasons for believing that 2+2=5, but despite this, is unable to do so. It's an example that neatly highlights the involuntarism of belief and the primacy of epistemic reasons.

The claim here isn't that we can't, in cases like Pascal's Wager, manipulate ourselves into indirectly believing something at will – I can change my belief that the lights are on simply by getting up and turning them off, for example – but this isn't what's at issue here. What's being denied is essentially that I can just directly will that *p* is true.

However, despite the appeal of the argument for doxastic involuntarism, it has been widely criticised both historically, by some such as Descartes (1996), William James (1999) and William Clifford (1999), as well as more recently, by among others James Montmarquet (1986), Roderick Chisholm (1968), Mark Heller (2000), Matthias Steup (2000), Bas van Fraassen (1984), Jonathan Bennett (1990), Shah (2002) and Carl Ginet (2001).

The arguments against doxastic involuntarism are varied and come from a number of different angles and engaging them is a thesis in and of itself. As a result, given the argument above in favour of doxastic involuntarism, what I'm about to stipulate is, I hope, only partially table thumping.

I take it that the arguments against doxastic involuntarism can be shown to be wrong and that, by and large, our beliefs are involuntary. Further, should I form a belief that *p* based on a non-epistemic reason, this non-epistemic reason is opaque to me. In other words, even if I form a belief as a result of a desire or fear (as is the case in motivated believing, discussed below) I do so without knowing that I have formed a belief non-epistemically. To call that believing at will is, as Scott-Kakures (1994) notes above, to misunderstand what is being claimed when it is said that we can't believe at will (Scott-Kakures, 1994:78).

In addition, our experiences tell us that to believe at will is impossible. (This is something that I believe will be more apparent when I examine Brian Ribeiro's (2002) discussion on scepticism below.) I certainly cannot think of an instance in which I've chosen to believe something. While I have believed things for all sorts of bad epistemic or non-epistemic reasons, those reasons were opaque to me at the time of

forming the belief. And further, having realised that the belief was formed or sustained by non-epistemic reasons those beliefs have ceased to exist.[36]

That said, I take that we are able to withhold forming a belief; and this we seem to be able to do voluntarily.

As John Cottingham (2002) points out, there is "a distinction … between positive and negative control over our beliefs …" (Cottingham, 2002:353). This distinction, drawn by Robert Audi (2001), defines negative doxastic control as "the ability to prevent their [beliefs] formation at will" (Audi, 2001:94). And "perhaps we have the latter but not the former, that is, one may hold that belief-formation is involuntary, while at the same time maintaining that we can deliberately prevent it or suppress it" (Cottingham, 2002:353).

This is something I'm happy to accept without thinking that it lessens the claim that I cannot directly form a belief at will.

I will not say much more about doxastic involuntarism in the rest of this project. As mentioned above, the reason I mention it at all is because if beliefs are things we can choose, then biased beliefs, motivated beliefs, beliefs based on inconclusive evidence and *idées fixe* admit to a simple explanation: we choose to believe them or to continue to believe them. However, given doxastic involuntarism of the sort spelt out above, the existence of these beliefs cries out for explanation.

3.8. Motivated believing

Motivated believing occurs when an agent takes a non-epistemic reason, or reasons, as a reason to believe something. In particular, the notion is that a desire for *p* or fear of not-*p*, rather than an epistemic reason for *p*, causes a belief that *p*.[37]

[36] Or have ceased to exist as a result of those reasons. We can of course believe that *p* based on a non-epistemic reason, but having come to see that our reason was non-epistemic, give up those reasons as reasons for *p*, but still believe that *p* as a result of other, epistemic reasons.
[37] We can of course imagine cases where the non-epistemic reason for a belief that *p* is a sharp blow to the head causing one to believe that *p*.

As Scott-Kakures (2000) explains:

> Wishful thinking and self-deception are instances of motivated believing. In the most straightforward sort of cases, a desire that my daughter be the most talented dancer in the company or, perhaps, the anxious fear that I have a life-threatening illness leads me to come to believe that my daughter is the most talented dancer in the company or, perhaps, to believe that various symptoms are really quite ambiguous, and, no doubt, nothing to be concerned about. (Scott-Kakures, 2000:348)

As Sarah Stroud (2006) notes, instances of motivated believing are commonplace.

> It is well known that people seem consistently to overrate themselves: almost everyone believes he is an above-average driver, for instance, and 94 percent of professors think they do better work than their average colleague. (Stroud, 2006:509)

While Scott-Kakures (2000) notes that:

> Wishful believing is a familiar, much evidenced, phenomenon ... Brown and Dutton (1995) note ... that "self-perceptions become more biased when the trait is highly desirable or undesirable" (*p*. 1290). They conclude that "on virtually every conceivable positively valued trait, the majority of people think they are better than others" (*ibid*.). (Scott-Kakures, 2000:348)

I will not say much more about these beliefs. I bring them up because an account of the truth-belief relationship needs to make sense of them.

What is important to note is that given what has been said about beliefs regulating for truth, what we should expect to see are these beliefs disappearing in the face of evidence that highlights the non-epistemic nature of these beliefs. By and large, this seems to happen.

As a personal confession, I form motivated beliefs all the time when I play Scrabble. Often, when wanting to put down a word that doesn't exist, I am convinced that I have seen this word before; that it exists. However, the moment I look in a dictionary my conviction that the word I have exists disappears. At this point, I am able to see that I formed a belief that *p* because I wanted *p* to be true.

Thus, while motivated beliefs are clearly an instance of epistemic error, they do tend to fit in with the truth-tracking features of belief we've explored above.

3.9. Biased beliefs

As Judith Baker (2011) and Stroud (2006) have argued, we don't always appear to form beliefs in line with the truth when it comes to forming beliefs about our friends or in trusting relationships. In these instances non-epistemic factors seemingly come into play, causing our commitment to the truth to become clouded.

As Stroud notes: "[friendship] places demands not just on our feelings or our motivations, but on our *beliefs* and our methods of forming beliefs" (Stroud, 2006:499). These non-epistemic factors interfere with our belief's ability to track the truth.

I take this as an obvious feature of belief. The anti-Semite, for example, is likely to form biased beliefs about Jewish people because his/her ability to track the truth is clouded by his/her anti-Semitism. As with motivated beliefs, above, I raise them as something an account of the truth-belief relationship needs to explain.

3.10. False beliefs

In addition to the fact that non-epistemic factors such as friendship, desires and fears can influence our beliefs, we get our beliefs wrong disturbingly often. Misunderstanding evidence, false testimony, an inadequate scientific framework, stupidity and so on are all common enough reasons for us to get it wrong.

However, while some of our beliefs are false is a trivial and obvious feature of belief to highlight, it is still something that any account of the truth-belief relationship must be able to account for.

3.11. Beliefs based on inconclusive evidence

Another common feature of belief is that "people form beliefs on the basis of inconclusive evidence all the time …" (Owens, 2002:393). Again, as with false beliefs, this feature of belief wouldn't be worth mentioning if it wasn't for the fact that I don't think that normative accounts can make sense of this feature.

As with false beliefs and biased beliefs, I'm going to simply stipulate this feature of belief. Though, it is worth mentioning that from the agent's perspective he/she must still take the believed proposition as True.[38]

3.12. *Idées fixe*

In addition to seemingly poor belief-forming practices, we don't always respond well to counter evidence. In fact we encounter people whose beliefs maintain even in the face of strenuous objection and counter evidence.

Erik Nisbet and Kelly Garrett (2010), for example, say that "more than 70 million Americans think that President Obama was born in a foreign country, despite extensive evidence to the contrary" (Nisbet & Garrett, 2010).

In this section I explore four different *idées fixe*: bigoted beliefs; delusional beliefs; beliefs that fail to shift in the face of unanswerable arguments; and beliefs that, as Van Inwagen (2009) puts it, "I 'was going to believe anyway'" (Van Inwagen, 2009:17).

The first of these *idées fixe* is common enough to my mind to require a brief explanation, but little defending. The second, delusions, are contentious. I offer a brief

[38] Jonathan Adler (2002) and Michael Hicks and Adler (2013) think that because an agent must take the believed proposition as True, he/she must also think that he/she has sufficient reasons for the Truth of that belief. However, as the *idée fixe* of Ribeiro (2002) below indicates, there may be cases when this isn't the case.

defence of these as genuine instances of belief. The third ones are extremely interesting and have ramifications for some of the earlier features of belief and need both a little more explanation as well as a little more defending. The fourth type of *idée fixe* seems to make sense intuitively, but is hard to defend other than to note that I think many people, including myself, have this type of belief. I leave these beliefs in because I believe that my account of the truth-belief relationship is able to explain these, while others struggle.

Before I move on, I take it that some cases of *idées fixe* are also cases of motivated believing. Thus, bigoted beliefs, for example, are probably (and I think I'm making a fairly reasonable assumption here) based on, or formed by, a fear of the other or a desire to be seen as better than the other.

3.12.1. Bigoted beliefs

The type of bigoted beliefs I am interested in are those racist, homophobic, sexist or otherwise bigoted beliefs that maintain despite being exposed to what appears to be unequivocal counter evidence.

Kwame Anthony Appiah (1990), in his discussion of Professor William Shockley (an educated and highly intelligent man with a Nobel Laureate in physics), notes:

> The objection to Professor Shockley's claims about the intelligence of blacks is not just that they are false; it is rather that Professor Shockley seems, like many people we call "racist," unable to see that the evidence does not support his factual claims and that the connection between his factual claims and his policy prescriptions involves a series of non sequiturs. (Appiah, 1990:209)

There seems little to defend here in terms of this feature of belief. These beliefs seem commonplace and I suspect everyone has met someone whose bigoted belief fails to change in the face of readily available counter evidence.

In addition, this type of belief tends to cohere with at least one other feature of belief listed above. The story an agent with this type of belief tells when confronted is an epistemically rationalising story; even if we don't think that explanation is a good epistemic story.

That said, and what requires explanation, is that these beliefs don't appear to regulate for truth. While the epistemically rationalising story an agent tells explains away the possible non-epistemic motivation, what cries out for explanation is the recalcitrance of this belief in the face of overwhelming counter evidence. The ability to explain this doxastic recalcitrance is lacking in the existing accounts of the truth-belief relationship.

3.12.2. Delusional beliefs

There is some argument about whether delusions count as beliefs. Some, like Robert Klee (2004), deny that delusions are beliefs at all. While others, such as Lisa Bortolotti and Matthew Broome (2008), have argued that they are.

The reason I keep these beliefs in, though I believe they represent a real challenge to my account, is that in some cases the delusion is from the agent's perspective indistinguishable from Vahid's (2006) concept of a belief. The agent takes the proposition as True for no extrinsic reason that is transparent to him/her other than the fact that he/she takes it as True.[39] To leave them out, therefore, seems to be an *ad hoc* move to protect myself.

In addition, I have argued above that what makes these types of delusions candidates for belief is that they share many of the features of belief – barring their resistance to what we take to be overwhelming counter evidence.

[39] With regards to those delusions that the person who has them doesn't think are True in a way that suggests a belief, I'm happy to accept that these are not delusions. Should an agent persistently hear voices in his/her head, but acknowledge that those voices don't "exist", I would claim that the agent has a delusion, but not a belief. The delusional mental propositional states I'm interested in are the MPS=Ts that despite being delusional are nevertheless ones that an agent takes as being True in a way that suggests a belief.

As Elsabe Jordaan (2012) explains:

> Delusions are beliefs that are firmly held contrary to reality and contrary to what most members of a society would regard as reality. Garety (1991) defines delusions as false personal beliefs that are firmly and consistently held despite logic or experiences to the contrary. (Jordaan, 2012:205)

I take note of them, as existing accounts do little to explain how delusions are compatible with their account of the relationship between beliefs and truth. Further, while some deny that these are beliefs, others, like myself, see those delusions where the agent takes the proposition as True in a way that suggests a belief as a delusional *belief*, and, therefore, in need of explanation.

3.12.3. Belief in the face of unanswerable, yet unbelievable arguments

Ribeiro (2002) details a case in which despite finding an argument unanswerable, thereby undermining all the epistemic reasons he has for the proposition, he nevertheless continues to hold onto the belief.

In this case, Ribeiro (2002) has strong evidence, via argument, to suspend his belief and yet seems not to be able to do so. As he confesses:

> I, BR, believe there *is* a cogent and intuitive argument for external world scepticism. I am quite convinced that this argument is unanswerable. Indeed, I have been convinced of this for quite some time. I have considered all of its facets and many of the objections which people will be inclined to make against it, and I feel satisfied in a deep and abiding sense that it is a good argument. Strangely though, this argument does not move me in the least. (Though it is a strange personal confession to make) I have tried, quite seriously, to allow myself to be moved by it. I have thought to myself, "BR, you yourself will say that you have no adequate grounds at all for believing, e.g., that you have a body. This belief of yours is not justifiable and this, as you commonly avow, is a serious objection in

> your book. You take the search for truth seriously and regard epistemic matters with great earnestness. Come now and be moved by what you take to be an unanswerable argument."
>
> But I am not moved. I still believe that I have a body and that my wife is real, that the floor before me really is there, and that I should act to insure the continued health of this apparent mortal coil. (Ribeiro, 2002:475)

In this instance Ribeiro's (2002) belief in an external world remains, despite his acknowledgement that he should, at the very least suspend his belief. Again, much like the bigoted belief above, what makes these beliefs interesting is that despite a strong link between belief and truth, indicated by the fact that beliefs tend to regulate reliably and unconsciously for truth, these beliefs maintain in the face of strong counter evidence.

This appears to be a genuine doxastic phenomenon, and therefore in need of explanation via the truth-belief relationship.

I find I have the same reaction to certain arguments. They appear to lead to an irrefutably true conclusion and yet, at the same time, I find them wholly unbelievable.[40]

[40] A personal favourite is one of Zeno's paradoxes of motion:

> Suppose a very fast runner—such as mythical Atalanta—needs to run for the bus. Clearly before she reaches the bus stop she must run half-way, as Aristotle says. There's no problem there; supposing a constant motion it will take her 1/2 the time to run half-way there and 1/2 the time to run the rest of the way. Now she must also run half-way to the half-way point—i.e., a 1/4 of the total distance—before she reaches the half-way point, but again she is left with a finite number of finite lengths to run, and plenty of time to do it. And before she reaches 1/4 of the way she must reach 1/2 of 1/4 = 1/8 of the way; and before that a 1/16; and so on. There is no problem at any finite point in this series, but what if the halving is carried out infinitely many times? The resulting series contains no first distance to run, for any possible first distance could be divided in half, and hence would not be first after all. However it does contain a final distance, namely 1/2 of the way; and a penultimate distance, 1/4 of the way; and a third to last distance, 1/8 of the way; and so on. Thus the series of distances that Atalanta is required to run is: …, then 1/16 of the way, then 1/8 of the way, then 1/4 of the way, and finally 1/2 of the way (of course we are not suggesting that she *stops* at the end of each segment and then starts running at the beginning of the next—we are thinking of her continuous run being composed of such parts). And now there is a problem, for this description of her run has her travelling an *infinite* number of *finite* distances, which, Zeno would have us conclude, must take an infinite time, which is to say it is never completed. And since the argument does not depend on the distance or who or what the mover is, it follows

One example is David Hume's (1977) argument against induction (Hume, 1977:15-20). I have gone through the reasoning. I am convinced he is right and yet despite feeling as though I ought to believe the conclusion that induction is unreliable I find I cannot.[41] A friend of mine, meanwhile, confesses to finding the Ontological Argument for God's existence equally unanswerable and yet unbelievable.[42]

> that no finite distance can ever be traveled, which is to say that all motion is impossible. (Huggett, 2010)

Now, I find this conclusion wholly unbelievable. I fundamentally believe in motion. I believe that when I take a step I'm covering a distance and that the motion is not illusory, but real. I think the laws of motion are scientifically valid and so on. Despite this, I find the argument above unanswerable. The argument structure is valid, the claims obviously true and so, it follows, the conclusion *must* be true. And, yet, despite having no answer to the argument my belief in motion is unchanged. In addition, I have no argument against Zeno other than a table-thumping "obviously motion exists". This is not to say that someone else doesn't have a response to Zeno's argument and, therefore, finds the argument answerable and thus unbelievable. What I'm highlighting is that there are cases in which some arguments appear to leave us thinking that our belief appears not to be based on a good-enough epistemic reason to maintain that belief. In addition to Zeno's paradox, I find arguments for external world scepticism enormously cogent and impossible to undermine and yet I find myself interacting with the world without hesitation, fully believing that it is real. As a result, I find G.E. Moore's (1939:147) common-sense response to external world scepticism as appealing as it is frustrating. It's appealing because the claim "Here's one hand, and here's another", therefore an external world exists, seems blindingly obvious to me. However, it fails, to my mind, to provide me with an adequate reason to either a) believe that the external world exists or b) believe that strong arguments for external-world scepticism are incorrect.

[41] It is worth noting that perhaps the reason that we can't believe the arguments as result of some feature of sceptical arguments rather than of the belief. Perhaps some arguments simply cannot cause the targeted beliefs to change. If this is the case then an account of the truth-belief relationship needn't worry about explaining these *idées fixe*; as the feature belongs to this type of argument rather than belief. In *Meditations on First Philosophy*, Descartes (1996) considers three sceptical scenarios (Descartes, 1996:12-15). The first is that since his senses sometimes deceive him, how does he know that they don't always deceive him. The second is that he has had dreams that he was unable to distinguish from reality, and might it not be the case that he is always dreaming, or, at the very least, dreaming now. The third is that perhaps an evil demon "of the utmost power and cunning may be deceiving him" (Descartes, 1996:15). In this final sceptical scenario Descartes suggests that the presence of an evil demon should cause him to question everything. One way of reading this final sceptical argument is to think that even mathematical truths and others of a similar ilk should be questioned. But, if this is the case, then the argument appears to be self-defeating. If mathematical truths are to be questioned, then logic should be questioned. If logic should be questioned, then truths based on logic should be questioned. If logical truths and logical arguments should be questioned, then any argument should be questioned. Since any argument that is based on logic should be questioned, then I have no reason to believe an argument the conclusion of which tells me that logic should be questioned. In this case if the scepticism tells us to give up logical truths it is seemingly self-defeating. If it's self-defeating then there is no reason to believe the conclusion. And so, if we, like Ribeiro find ourselves not believing the conclusion, then it's not as a result of some feature of belief that needs to be explained, it's as a result of a flaw in the logic of the argument. Now, what I've said above needs a lot more defending than I've given it here. However, the point isn't to defend the above argument, but merely to raise it. If it is a good argument, then we have a reason to think that the reason why we don't/can't believe sceptical arguments is because of the argument, rather than because there is a feature of belief that needs to be explained. While I think there is some merit to the argument above, I still think there is a feature of belief to be explained here. It seems to me that sceptical arguments can attack either a proposition (motion is possible) or a method of reasoning (such as Hume's (1977) argument against induction (Hume, 1977:15-20)). Now, it seems to me that should a methodological

As with the other *idées fixe* above we have a belief that simply fails to regulate for truth and this requires a model of the truth-belief relationship that explains how this is possible.

Before moving on to the final type of *idée fixe*, it is worth making note of one thing.

Ribeiro's (2002) confession and this evidence-resistant phenomenon appear to bolster the case for doxastic involuntarism. Ribeiro thinks his belief should change. In fact, he has "tried, quite seriously, to" (Ribeiro, 2002:475) change his belief, because he doesn't think he has a good reason to maintain his belief. However, despite this, despite trying to, at the very least, suspend his belief, he still has the belief. While the fact that this belief maintains is the focus of this section, Ribeiro's confession also highlights the involuntarism of belief.

---

scepticism attack the very method being used to make the sceptical claim that argument is self-defeating and it is self-defeating because the argument tries to undermine the very method it's using to undermine the method. If all sceptical arguments are like this, then we should not be surprised to find that our beliefs fail to be moved by them. The problem is with the argument, not the belief. However, two things count against this position and suggest that at least some of the time this type of *idée fixe* is in need of an explanation by the truth-belief relationship. The first is that while some sceptical arguments may be self-defeating not all sceptical arguments are methodologically self defeating. In the case of Zeno's paradox in the previous footnote, the sceptical argument attacks a claim rather than a method of reasoning. So, if there's nothing wrong with the argument, then there's something about the belief that needs to be explained. The second is that even if there was something wrong with all sceptical arguments, the agent would have to appreciate that in order for us to look to the argument rather than the belief as the explanandum. Let me explain. In much the same way that false testimony can lead to a belief, bad arguments, invalid arguments, unsound arguments, all cause beliefs. If an agent doesn't think that there is something inherently wrong with all sceptical arguments (I don't, for example) then what we should expect to see is an agent's belief change in the face of the sceptical argument. So, even if there is something wrong with all sceptical arguments this would have to be something that the agent appreciates. Otherwise he/she would simply treat sceptical arguments as any other type of argument. Given that other types of arguments can lead to believable conclusions, and given an agent's failure to see sceptical arguments as inherently unreliable, there is no reason for a sceptical argument to fail to form a belief. Why is it then that some arguments while admitting to no answer also cause no belief? The answer, it seems, is that we have a type of evidence resistance that applies to these beliefs. And this is a feature of belief that needs to be explained.

[42] It is important to note that in cases where one finds the argument (sceptical or otherwise) answerable and unbelievable there is nothing to explain. In this case your belief hasn't changed because you have no reason to change your belief. You have what you take as good epistemic reasons to maintain your belief. In the case of an unanswerable yet unbelievable argument, however, you have good reasons, epistemic reasons, to, at the very least, give up your belief. And yet, despite this acknowledgement you find yourself still with the belief in question.

3.12.4. Beliefs that I would have believed anyway

The final type of *idée fixe* I would like to consider comes from a passage from Van Inwagen's (1993) *Metaphysics*.

In his introduction he comments:

> Nevertheless, I have to admit that there is probably some truth in the charge that there are certain theses that I "was going to believe anyway" – theses that I would believe no matter what philosophical arguments I was presented with. (Van Inwagen, 1993:15)[43]

In this case it's not that we find an argument "*unanswerable* and yet *unbelievable*" (Ribeiro, 2002:468), it's that we know that irrespective of the argument put forward, or despite evidence to the contrary, we're going to carry on believing something anyway.

For example, I have stopped being able to trot out John Corvino's (1997) excellent defence of same-sex relationships. I know that I'm morally comfortable with same-sex relationships and know that whatever argument is presented against this belief I will probably not engage it on its merits. Instead I will simply hold onto my belief. My belief that same-sex relationships are morally acceptable is something that I'm almost certain "I would believe no matter what philosophical arguments I was presented with" (Van Inwagen, 1993:15).

[43] Interestingly, in the third edition of this book the phrase "... theses that I *would* believe no matter what philosophical arguments I was presented with" (Van Inwagen, 1993:15) (italics added), is "... theses that I *should* believe no matter what philosophical arguments I was presented with" (Van Inwagen, 2009:17) (italics added). Both of these are interesting claims, implying that we can and/or do sometimes hold onto beliefs even when the evidence for the belief is inconclusive or the veracity of the belief has been strongly challenged. One way of reading the later formulation is that we have certain beliefs that we should believe irrespective of the evidence against them. Thus, it might be prudent always to believe that one is an agent, even if all the evidence suggests that one is actually a determined being. Or, alternatively, we should always believe in the equality of rational agents even if the evidence suggests that the equality is nothing more than philanthropic naivety. However, the implication with the later formulation is that we might not be able to believe these things even though maybe we ought to, in some moral or pragmatic sense. Van Inwagen's earlier formulation is far more revealing of the type of doxastic scenario that I'm interested in.

Now there could be many epistemically valid reasons for this. I might, for example, believe that the arguments in favour of the moral acceptability of same-sex relationships are so convincing that it's unlikely that any counter argument could ever be epistemically valid. That said, I can't imagine any counter argument being something that would sway my belief and, given the strong link between belief and truth, it seems odd that I could hold onto these types of belief.

Again, like the other three examples, we appear to have beliefs that maintain in the face of counter evidence.

Given the legitimacy of these evidence-resistant beliefs as doxastic phenomena our account of the truth-belief relationship must be able to explain them.

An account of the truth-belief relationship, therefore, needs to do two seemingly inconsistent things. One, it needs to be sufficiently strong to make sense of those features of belief that indicate a strong link between beliefs and truth (3.1 – 3.7). Two, it needs to be sufficiently weak/flexible enough to account for beliefs that seemingly don't track the truth (3.8 – 3.11).

Any account that cannot make sense of beliefs regulating for truth, for example, fails to capture what it is to believe. Similarly, any account that cannot explain doxastic error equally simply fails to capture what it is to believe.

In the following section I make explicit an argument in the literature that an adequate account of the truth-belief relationship must be able to explain the features of belief.[44]

[44] Given the existence of *idées fixe* and motivated beliefs there is one potential objection to the notion of belief that I put forward earlier. Above I argued that Vahid's (2006) notion of belief is that to believe that *p* is to take *p* as True not for any extraneous reason transparent to the agent, but simply because the agent takes *p* as being True. However, given the existence of motivated beliefs and *idées fixe* surely these are clear examples in which someone believes that *p* not because he/she takes *p* to be true, *tout court*, but instead takes *p* to be True in order to fulfil some other goal. If this is right, then Vahid's concept is in trouble. However, the important part of the concept of belief that I have chosen is the caveat that I added above, namely that if an agent forms a belief for an extrinsic reason that reason is opaque to him/her. What allows us to call these mental propositional states beliefs is that when one forms an imagining, assumption or guess that *p* one does so in the knowledge (and I use the word "knowledge" colloquially here) that one is taking *p* as True for the sake of something else. When one forms a motivated belief or has an *idée fixe*, on the other hand, one is unable to think that this mental state is thought of as True for any other reason or purpose other than that one takes *p* as being True. As Ward Jones (2002:217) says:

4. The importance of the relationship between belief and truth

Embedded in the literature is an argument, though not explicitly made, that an account of truth-belief relationship must, in virtue of that account, be able to explain the features of belief. There are several claims in the literature that support this.

Timothy Chan (2013), discussing doxastic involuntarism, notes:

> Rather than taking belief aiming at truth as a datum which constrains our ability (or inability) to believe at will, the current point takes doxastic involuntarism itself as a datum, which constrains the account one can give of belief's aiming at truth. (Chan, 2013:7)

Owens (2003) observes that ascribing a goal to beliefs explains why beliefs regulate for truth.

> Postulating a goal for belief explains not only the authority of epistemic norms but also their motivational efficacy. And, in the eyes of many philosophers, this is essential for epistemic norms to have any authority. (Owens, 2003:284)

Both of these claims indicate that what makes an account of the truth-belief relationship successful is its ability to make sense of the features of belief.

> It has often been claimed that our believing some proposition is dependent upon our not being committed to a non-epistemic explanation of why we believe that proposition. Very roughly, I cannot believe that *p* and also accept a non-epistemic explanation of my believing that *p*. (Jones, 2002: 217)

Now, when I'm in the grips of a motivated belief I am, at the time of having that belief, convinced of its veracity for its own sake. Should someone say "you only believe that because you want to [fill in some non-epistemic goal]", I would sincerely deny it, or in accepting my interlocutor's claim, give up my belief. The moment I become aware of the fact that my belief is based on some desire or fear rather than evidence in order to serve some other purpose than truth, my belief disappears. As noted above, this is true even in the case of delusional beliefs. Should an agent see that his/her belief is based on some non-epistemic reason the belief (though perhaps not the delusion) disappears. What makes delusions interesting is just how hard it is to get someone with a delusion to see that his/her belief is based on non-epistemic reasons. Further, even in the case of Ribeiro's (2002) sceptical challenge, while the inference to the best explanation is that our belief is based on a non-epistemic reason the agent is unable to say what that reason is. The agent thinks that he/she lives in the "real" world, but can't say why. This, however, does not amount to the same claim that I believe that *p* is true for some extrinsic reason. Thus, while Vahid's (2006) claim needs a little finessing, it is still, I think, essentially correct.

Shah and Velleman (2005), meanwhile, argue that one of the problems with the teleological account is that it can't make sense of transparency. There are two claims about the responsibilities of an account of the truth-belief relationship tacitly being made here. The first is that an adequate account of the truth-belief relationship must, in virtue of that account, be able to explain the features of belief. The second is that any account that fails to explain these features of belief, fails as an adequate account of the truth-belief relationship.

Both claims seem right to me. What makes normative accounts appealing, argue Shah and Velleman (2005), is the fact that they are able to explain features of belief that teleological accounts can't. And while I'm not sure that normative accounts *are* better than teleological accounts, I do think it is correct to claim that the worth of an account of the truth-belief relationship lies in its explanatory power.

Being able to explain a feature of belief suggests that the account that provides that explanation is the right way to understand the relationship between beliefs and truth. In addition, failure to account for a feature of belief counts against a theory of the truth-belief relationship. If theory *t* of the truth-belief relationship unequivocally rules out the possibility of feature *f*, then given *f* either we must scrap *t* in the face of *f* or we must be able to show that *f* isn't a feature of belief and is, therefore, not something that *t* needs to account for.

5. Conclusion

The aim of this chapter has been to lay out several features of belief. In cases where I felt the feature was sufficiently intuitive, or already part of the extant literature, I've explained it and moved on. In some cases, where I feel that the feature is contentious and, therefore, might threaten some of my arguments later on I've argued for the feature in question.

I then made explicit an argument in the literature that the job of any adequate theory of the truth-belief relationship is to be able to account for the features of belief listed above.

There are two prevalent models found in the literature of how to understand the truth-belief relationship: either teleologically or normatively.

In the following two chapters I argue that both accounts fail before turning to my own account.

Chapter 3: The teleological account

1. Introduction

In the previous chapter I noted and argued for several features of belief. Those features are:

1. The relationship between belief and truth is direct.
2. Because the direct link is between belief and what an agent takes as True, rather than the Truth, the scope of belief is vast and varied. An agent A can believe that *p* at the same time as another agent B believes that not-*p*.
3. Beliefs can be both occurrent and dispositional.
4. Only what an agent takes as an epistemic reason counts for him/her as a reason for belief.
    a. As a corollary, unless you're in the face of something like Brian Ribeiro's (2002) doxastic crisis, when you have a belief that *p*, you find that you tell an epistemic story about that belief.
5. Epistemic reasons have a direct link with belief, such that if you find epistemic reason *r* for *p* compelling, *r* would directly cause you to believe that *p*, without the need to weigh *r* against instrumental reasons for or against *p*.
6. Beliefs regulate for truth.
7. Beliefs are involuntary.
8. We form motivated beliefs.
9. Some of our beliefs are biased.
10. We form false beliefs.
11. We form beliefs based on inconclusive evidence.

And, finally,

12. Some beliefs are *idées fixe*.

Having laid out these features of belief, I highlighted an argument in the literature that any account of the truth-belief relationship must be able to explain, in virtue of that relationship, the features of belief. I argued that the credibility of an account that cannot explain a feature of belief is undermined as a full account of the truth-belief

relationship. An adequate account of the truth-belief relationship, therefore, needs to explain not only when we get it right, but also when we don't.

Perhaps the best-known explanation of the truth-belief relationship is Bernard Williams' (1973) potentially illuminating phrase "beliefs aim at truth" (Williams, 1973:136). While Williams (1973) may not have meant this metaphor to be taken literally, people such as David Velleman (2000), Conor McHugh (2012) and Asbjørn Steglich-Petersen (2009), among others, have taken it more-or-less literally. They argue that the way to understand the aim-at-truth metaphor and, subsequently, the relationship between belief and truth, is teleologically.

In this chapter, I explore the viability of teleological accounts as a version of the truth-belief relationship.

I begin by arguing in Section 2 that there is a crucial distinction missing from the literature between two different ways of understanding teleology: namely, either as an aim or as a purpose. The failure to recognise this distinction means that expositions of teleological accounts, as well as the objections to teleological accounts, are muddied. Addressing this distinction, spelling out each of the accounts individually, means being able to see the relative strengths and weaknesses of each, as well as more accurately understanding the objections in the literature to teleological accounts.

Having defended this distinction, in Section 3 I lay out my version of an Aim Account. Aim Accounts see beliefs intentionally aiming at truth in much the same way that an agent can aim to get up early to watch the sunrise.

In Section 4 I lay out my version of a Purpose Account. These accounts see truth as the purpose of belief in much the same way that kettles have the purpose of boiling water.

Having laid out these two separate accounts, in Sections 5 and 6 I consider two objections found in the literature that have been levelled at teleological accounts.

The first is an objection by David Owens (2003), who argues that given what it is to be an aim, beliefs cannot be said to aim at truth. Given that the distinction I've drawn between the two types of teleological accounts is a new one, I examine the ability of each of these accounts to deal with this objection separately. I argue that while Aim Accounts are threatened by Owens' objection, Purpose Accounts can comfortably deal with it.

The second objection is Nishi Shah's (2003) "Teleologist's Dilemma" (Shah, 2003:460). Shah argues that teleological accounts must either give up their ability to explain transparency, or must give up their ability to explain instances of motivated believing, either of which represents a fatal flaw for these accounts (Shah, 2003:460-465). In this instance I argue that it is Purpose Accounts that struggle to overcome Shah's objection, while Aim Accounts are able to meet the challenge.

The fact that Aim Accounts and Purpose Accounts respond differently to the challenges set out in Sections 5 and 6 suggests that the distinction made in Section 2 is a legitimate and important one.

Having examined these two objections, in Section 7 I consider the broader question of whether these accounts, Purpose or Aim, can make sense of the features of belief detailed in Chapter 2. I argue that these accounts, no matter how they are formulated, fail to capture several features of belief. I conclude, therefore, that while teleological accounts are misunderstood they nevertheless fail as accounts of the truth-belief relationship.

Before moving on to the next section, it is worth reiterating that what we are exploring is how to make sense of the truth-belief relationship. This is a claim about the relationship between beliefs and truth, rather than between believers and truth.

It seems obvious to me (as discussed in greater detail in the previous chapter) that I, as a believer, can have a variety of goals in terms of forming a belief. I can aim at forming a true belief by intensely researching a topic and refusing to give credence to the proposition under consideration until I'm utterly convinced of its truth. However, I can just as easily aim at forming a false belief, going to a hypnotist, engaging in

epistemically dubious practices with the intention of undermining my epistemic virtues and so on. However, the epistemic capriciousness of doxastic agents doesn't undermine the possibility that beliefs *qua* beliefs aim at truth; however "aim" is to be understood. There is a difference between a believer aiming at truth and a belief aiming at truth.

As Ward Jones (2002) puts it:

> Believing is often a goal-oriented process, but it is not (always) so in virtue of desires or intentions that the believer has. In the same way, the maintenance of my body temperature is a goal-oriented process, but it is not so in virtue of desires or intentions that I have. (Jones, 2002:222)

What we're exploring in this chapter is a claim about the relationship between *beliefs* and truth.

2. Aims and purposes

The terms "aim", "goal", "purpose", and so on, are used interchangeably in the literature. But to my mind there is a clear distinction between aiming at something and something having a goal or purpose.

To aim at something is to have the intention to achieve a certain end. I aim at (have the intention of hitting) a bulls-eye while playing darts, for example. This intentional pursuit of a target is what it means to aim. In terms of being an aim, it seems that each aim must have a target. It seems impossible to say of someone or something that "he/she/it is aiming at …" without providing an object in that sentence.

To have a goal or purpose, however, needn't mean having any intention. Toasters, for example, have the purpose of making toast. But toasters don't aim at making toast. It is simply the purpose of that device. Mechanisms and processes behave in such a way so as to fulfil the purpose of the toaster. To talk about intentions here would just be anthropomorphic.

It may also be the case, as St Thomas Aquinas suggests, that human nature is to be understood as having the purpose of reconnecting with the divine, or, as Aristotle thought, of living the Good Life. In this case the purpose of the agent is something that he/she *has*, rather than something he/she goes about aiming to achieve – though of course we may think that he/she should aim at it as well. Even if the agent never aims in the intentional sense at fulfilling these purposes, he/she can nevertheless be said to have a purpose.[45]

Thus, there seems to be clear distinction between the two. In the case of aims there is an intentional project to achieve some end. In the case of a purpose, it is an inherent end that draws, as it were, the agent or thing towards that end. In the case of the latter we have a purpose without any aim/intention to reach it, whereas in the former we have the aim to reach the goal in question. Thus, while we may not be able to talk about aims without targets or goals, we can talk about purposes without aims.

Williams' (1973) phrase that "beliefs aim at truth" (Williams, 1973:136), therefore, can be understood, teleologically, in two ways: either as beliefs aiming at something or as beliefs having truth as their purpose (aim). Therefore, it seems clear to me that there are two ways we can make sense of teleological accounts: what I will call Aim Accounts, which claim that beliefs intentionally *aim* at the target of truth, and what I call Purpose Accounts, where truth is the *purpose* of beliefs.

I believe this missing distinction has created two problems.[46] One, by conflating aims with purposes our understanding of what a teleological model can offer the literature

[45] There may be a concern that while purposes don't speak of having a direct aim, things like toasters are designed to have a certain purpose and can therefore be said to have an aim. Thus, someone builds a toaster with the intentional aim of it achieving a certain goal. I'm wary of this deferred intention. While the toaster-maker may have aims, it certainly doesn't seem to me that the toaster has aims. However, even if we do want to claim that toasters have aims in virtue of the artisan having had an intention for that toaster, I take it that some our systems, born out of a process of evolution, can legitimately be said to have a purpose or goal without there being any talk of intention. The homeostatic system, for example, can be said to have the goal of regulating our body temperature without any direct or indirect talk of intention creeping in.

[46] I have not seen this distinction elsewhere in the literature. From what I have seen, the person who gets the closest to making this distinction is Anna Edmonds (2007). She draws a distinction between the aims of the believer (what she calls Purpose Accounts) and the goals of belief (what she calls Teleological Accounts) (Edmonds, 2007:15-22). Her distinction, however, is between what we can say of the aims of the agent as they pertain to belief and what we can say of beliefs themselves in terms of

on the truth-belief-relationship has been undermined. Conflating the two means muddied and, to my mind, often incoherent expositions of teleological understandings of the truth-belief relationship. Two, the knock-on effect, of course, is that criticisms of teleological accounts are equally damaged by the failure to distinguish between the two.

Owens (2003), for example, in his discussion of the aim-at-truth metaphor, talks about belief as having "a goal which it is striving towards, or a function which it is seeking to discharge" (Owens, 2003:283). Talk of "striving towards" and "seeking to discharge" (*ibid*.) suggests that Owens thinks that the way teleological accounts should be understood is as what I'm calling Aim Accounts. Though, given that purposes would pull the processes towards that goal this talk of "seeking to discharge" (*ibid*.) could be understood as part of a discussion about purposes.

On the other hand, McHugh (2012), whose paper is a response to Steglich-Petersen's (2009) paper, which in turn is a response to Owens' (2003) paper, appears to see truth as the purpose of belief. McHugh, commenting on teleological accounts, notes: "According to the teleological account of belief, the attitude of belief has a characteristic and constitutive end, aim, goal or purpose" (McHugh, 2012:426).

This talk of belief having "a characteristic and constitutive end, aim, goal or purpose" (*ibid*.) suggests that McHugh (2012) seems to think that teleological accounts are best understood as what I'm calling Purpose Accounts. McHugh's response, therefore, appears to defend a version of the teleological account from Owens (2003) that Owens may in fact not be attacking.

Further, it's hard to know which camp proponents and detractors of a teleological account sit in. Certainly Owens' (2003) objections seem primarily levelled at an Aim Account. In objecting to the claim that beliefs aim at truth, Owens draws an analogy with guessing (Owens, 2003:289-300). Given that guessing is definitely an intentional *aiming* at truth, Owens' objection appears to rest on the claim that teleological accounts are Aim Accounts.

---

their aims/goals. While this distinction may ultimately and tacitly contain the same distinction that I'm drawing between aims and purposes, this distinction is not drawn out in Edmonds' work.

That said, Owens (2003) does talk about the purpose of belief as well (Owens, 2003:284-289). Whether Owens means this merely as a synonym for aim, or because, in aiming at truth, beliefs have the target of truth, or whether Owens is conflating Aim Accounts with Purpose Accounts is unclear.

This missing distinction means that the potential subtleties and nuances specific to these different accounts are lost. This undermines both the attempt to understand the truth-belief relationship teleologically as well as the objections to these types of accounts. Below I sketch up my versions – and to my knowledge, the first versions – of an Aim Account and a Purpose Account, respectively.[47]

3. Aim Accounts

In trying to make sense of the aim-at-truth metaphor, Ralph Wedgwood (2002) starts with the thought:

> It is often claimed that beliefs aim at the truth. Indeed, this claim has often been thought to express an essential or constitutive feature of belief. But this claim is obviously not literally true. Beliefs are not little archers armed with little bows and arrows: they do not literally "aim" at anything. (Wedgwood, 2002:268)[48]

Aim Accounts claim that beliefs do precisely what Wedgwood (2002) denies, namely, aim (in an intentional, targeted way) at truth.[49]

---

[47] It is worth noting one thing that neither Aim Accounts nor Purpose Accounts are committed to. The aim or purpose of Truth does not require us to search out and believe all available truths. As David Owens (2003) states, "the aim of truth [must be] stated negatively as the purpose of avoiding error; it imposes no obligation on believers to seek out all, or even some truths" (Owens, 2003:289). Thus the aim of truth is permissive rather than injunctive. The reason for this is that an aim of truth or the purpose of truth would apply specifically to one belief rather than all potential beliefs. Thus if my belief that *p* aims at truth the obligation in terms of the aim pertains only to that belief. Similarly, if the purpose of my belief that *p* is truth, then this pertains only to the obligations that a purpose imposes on that belief, rather than to anything else.

[48] Ralph Wedgwood (2002) does not draw a distinction between something like Aim Accounts and something like Purpose Accounts. He claims that since it can't be true that beliefs aim at truth in this intentional way, we must understand the aim-at-truth metaphor differently. His option is to see the claim as a normative claim.

[49] There is a metaphysical concern that I'm ignoring here. The Aim Account relies on the fact that beliefs intentionally aim at the target of Truth. What does the explanatory work in Aim Accounts is that

We need to be a little bit careful here. In Chapter 2 I drew a distinction between truth and Truth. The former is what an agent takes as the Truth, the latter is the objective Truth (whatever that might be). This distinction means that there are potentially two different types of Aim Account: those that postulate beliefs aiming at truth, and those that postulate beliefs aiming at Truth. However, to my mind, an Aim Account that sees beliefs aim at truth quickly rules itself out as a viable account of the truth-belief relationship.

As I noted in Chapter 2, what makes the relationship between belief and truth so interesting is the – for the most part, and beyond the control of the agent – irresistible motivational efficacy of epistemic reasons. Beliefs regulate for truth, even if we don't want them to and do so automatically, often below the level of full consciousness of the agent. But if beliefs aim at truth, rather than Truth, I don't think this account could explain why beliefs regulate for truth.

Let's assume that an Aim Account is correct and that beliefs aim at truth. Under this account, to believe that *p* is thus to take *p* as True in virtue of the aiming mechanism of beliefs. Let's now assume that there is evidence that not-*p* is True. Why would the agent's belief change? As far as the belief is concerned the target has been hit. If this

---

in virtue of aiming at the target of Truth beliefs are required to do certain things (follow certain epistemic norms) in order to hit this target. Now, if this is right, then in order to aim, there needs to be someone or something that intentionally does the aiming. While we can make sense of processes occurring, literally, without thought in the case of purposes (think toasters again), in the case of aims there is an intentional targeting of some goal. Forming an intention to hit something requires something like agency. If this is right, then an Aim Account appears to be committed to claiming either that beliefs are agents, which aim at Truth, or beliefs are sub-intentional agencies that aim at Truth. Of course, believers, as agents, can aim at Truth. Doxastic deliberation is precisely this type of project. But Aim Accounts, as spelt out in this section, describe the relationship between beliefs and truth rather than between believers and truth and what this means is that the belief must be the one with the intention. Supposing that beliefs are agents or sub-intentional agencies that drive at Truth requires us to make an ontological commitment that I worry raises more questions that it answers. To begin with, anything with agential powers it seems must be able to choose. Given that beliefs patently cannot choose, I wonder if the explanation of the truth-belief relationship in terms of sub-intentional agencies makes any sense conceptually (this appears to be part of Owens (2003) worry). It's a worry that I have to put aside here. I think that Aim Accounts fail for reasons other than the metaphysical worry. That said, while I think it's potentially problematic to postulate the existence of sub-intentional agents for beliefs, others suggest that this move may have merit. Ward Jones (2004), for example, thinks that explanations of "pragmatic believing is best understood as [beliefs being part of] a sub-intentional process" (Jones, 2004:3). Meanwhile, Mark Johnston (1988) thinks that self-deception is best understood as a result of sub-intentional agents. If we can make sense of self-deception in terms of sub-intentional agencies then the jump to see beliefs as sub-intentional agencies may not be as big as I'm concerned about.

is the case, if beliefs aim at what an agent already thinks is True, then it is hard to see how it could be that beliefs regulate for truth. To regulate for truth, it seems there must be a gap between truth and what the evidence suggests is True. If a belief aims at what it has already hit, there's nothing left to aim at.

If beliefs aim at Truth, however, then we can understand how the belief would regulate for truth. By aiming at Truth, the belief is required to shift the target in light of the new evidence. Evidence of the Truth of a claim, other than the one that the agent believes, creates a new target to hit for a belief that aims at Truth and so, if we assume that beliefs aim at Truth, we should expect the belief to change.

The reason for this is relatively clear. To have a constitutive aim of Truth, means that a belief always has the objective of hitting that target. If there is sufficient evidence for the Truth of *p*, then aiming at Truth would require forming a belief that *p*. Failure to follow the evidence would mean that the belief could no longer be said to follow its constitutive aim and so could no longer be said to be a belief. Thus, if beliefs aim at Truth, then we should expect them, in virtue of this constitutive aim, to change in the face of sufficient counter evidence. Thus, it seems that if Aim Accounts are to get off the ground, then they must aim at Truth rather than at truth.

Let's take it then that if we're talking about Aim Accounts we're talking about Aim Accounts that have Truth as their constitutive aim.

What makes Aim Accounts attractive is that they are able to explain several features of belief in virtue of the constitutive aim of belief. If a belief aims at Truth, we cannot take non-epistemic reasons as reasons for belief. To do so would be in direct conflict with the aim of belief. For if I take reasons other than epistemic reasons as reasons to believe, then I cannot be said to be aiming at Truth. Similarly, when we see that our reasons for our belief that *p* are insufficient for the truth of *p* we give up that belief. Should we come to see that our belief that *p* is based on or sustained by some non-epistemic reason we give up the belief. To fail to do either would be for beliefs to fail to aim at Truth.

So, beliefs respond to epistemic rather than prudential reasons because beliefs aim at Truth. To respond to prudential, rather than epistemic reasons, would be to give up on aiming at Truth. Thus, we are able to explain both the primacy and the motivational force that epistemic reasons, and only epistemic reasons, have on our beliefs. The only reasons that help beliefs achieve the constitutive goal of Truth are epistemic and if it is constitutive of belief that they aim at Truth, then beliefs must respond to epistemic, and only epistemic, reasons. Similarly, because our beliefs aim at Truth, our beliefs must respond to counter evidence, which accounts for beliefs regulating for truth.

A constitutive aim of Truth also explains why it's impossible to believe at will, as well as why we struggle to believe certain things about our beliefs. If, for example, we believe that our belief that *p* is false we are saying our belief no longer aims at Truth and, in so doing, rule it out as belief. For if beliefs are constituted by the aim of Truth and we believe that our belief no longer aims at Truth, it is now no longer a belief.

Furthermore, because it is the belief rather than the believer that is aiming at Truth we can make sense of beliefs forming and revising below the level of full consciousness as all the work is being done by the teleological relationship between beliefs and Truth: our unconscious beliefs are required to adhere to these epistemic norms, irrespective of our projects as believers, as a result of the constitutive aim of belief.

As it stands, an Aim Account is doing a good job. It explains several features of belief in a coherent fashion as a result of a constitutive aim of belief. In addition it is able to explain how our beliefs appear to follow epistemic norms below the level of full consciousness. Thus, what makes the Aim Account attractive, to use the words of Steglich-Petersen (2009), is that it "promises … a simple, unified and *prima facie* unproblematic explanation of doxastic normativity" (Steglich-Petersen, 2009:396).

4. Purpose Accounts

Unlike Aim Accounts, which suppose beliefs intentionally aiming at Truth, Purpose Accounts suggest that the inherent purpose of beliefs is truth. In much the same way

that toasters have the end-purpose of making toast (or of at least heating bread from somewhere between cold to coal), so too do beliefs have the purpose of truth.

As with Aim Accounts above, it's important to determine what we mean by truth here.

I take it that if Purpose Accounts are to explain the primacy and motivational force of epistemic reasons the purpose of beliefs must be Truth, rather than truth.

Imagine our beliefs as a specialised trap designed to catch only a certain type of animal, in this instance, truth. As our propositional content runs over the trap it is examined. Should one of these propositional claims meet the condition of truth, the trap springs shut and a belief is formed. Should evidence come to light that the belief is incorrect it is hard to see how, given this model, the belief would change. As far as the trap is concerned the right thing has been caught and so there seems no reason to release what it has. What could possibly account for the trap releasing the "incorrect" belief in favour of the "correct" belief?[50]

There are two potential but ultimately unsuccessful answers to this question. One, the trap resets itself. There are two difficulties here. The first is that if we think the trap resets itself we must ask why it resets itself. As far as the trap is concerned the right thing has been caught and so there is no reason to reset. The second is to suggest that it's not the trap that resets itself, but rather some other mental construct (let's call this a hunter) resets the trap. But if a hunter resets the trap, then it is not beliefs that have the purpose of truth, but the hunter. If this is the case, then what does the explanatory work is not the truth-belief relationship but the truth-hunter relationship. Further, if the hunter is resetting traps to release previously-thought-of-as-True mental states in order to catch now-think-of-as-True mental states this suggests that Truth, rather than truth, is the purpose of the hunter. If a hunter is convinced that he/she has caught is what he/she wanted there is no reason to reload the trap. If, however, the hunter has the purpose of Truth and sees that what he/she has caught is not the Truth, but something else, then he/she has a reason to reset the trap.

[50] I use the quotation marks here as a belief that an agent gives up may well be True, while the belief that an agent forms in favour of the one he/she has given up may well be false.

Two, it's not that a truth is caught and released in cases of beliefs regulating for truth, instead a new belief is formed (a new truth is caught), which takes precedence over the old one. However, while we might appeal to a mechanism that allows us to forget the previously-thought-of-as-True mental state in favour of the now-think-of-as-True mental state there seems to be a) no reason to think that we have an infallible mechanism causing us to forget the former in favour of the latter and b) no reason to think that this would happen. Why would we forget what we previously-thought-of-as-True the moment something new was caught? If belief's purpose is truth, then from this perspective the former mental state is as veridical as the latter mental state.

If the purpose of beliefs is truth, then once a belief is formed there seems no reason for this belief to change in the face of damning counter evidence. If beliefs can't change in the face of damning counter evidence, then an account with truth as the purpose of belief cannot explain the motivational force of epistemic reasons. Since epistemic reasons have this motivational force, causing our beliefs to regulate for truth, an account that suggests truth as the purpose of belief cannot be the proper account of the truth-belief relationship.

If, however, a belief's purpose is Truth, then we should expect to see beliefs becoming "truer", as it were, in the light of evidence for a proposition.

A doxastic trap that only shuts when something is True, could explain why there is still work that can be done when a belief encounters strong counter evidence to the contrary.[51] Failure for the belief to react in the face of strong counter evidence would mean a failure to fulfil the inherent purpose of belief.

This type of Purpose Account would also be able to explain why beliefs follow epistemic norms. To have a purpose is inherently to have a goal that the thing works towards. As a result of this purpose there are certain inherent conditions that

[51] There is a very valid concern that I spell out in greater detail below about how this trap would shut in the first place. As the model stands, a belief could only be formed once the mental was actually True. Since we have false beliefs this is a problem for this type of Purpose Account. For me, this is a damning objection. However, for the moment all I want to show is that a Purpose Account with Truth as its purpose could explain why beliefs regulate for truth as well as make sense of a number of epistemic norms, while one with truth as its purpose could not.

need/must/ought to be met in order to achieve the purpose. For beliefs to meet their purpose of Truth they would need to follow certain epistemic norms. In this case, only epistemic reasons can count as reasons for a belief. Taking on board non-epistemic reasons would mean the purpose of belief is something other than Truth. For much the same reason beliefs based on or sustained by non-epistemic reasons must disappear: seeing one's belief as being based on or sustained by non-epistemic reasons means to have failed to meet the inherent purpose of belief.

Ascribing the purpose of Truth to the belief, rather than the believer, explains why these norms are inherent, rather than forced upon the mental state by the believer, and this explains why so much of the adherence to doxastic norms occurs below the level of full consciousness.

As Owens (2003) points out:

> Postulating a goal for belief explains not only the authority of epistemic norms but also their motivational efficacy … the explanatory strategy behind it is perfectly sound: if believing were purposive, we could explain the above features of belief by reference to belief's purpose. (2003:284-289)

As it stands, both Aim Accounts and Purpose Accounts seem in good shape. Both are able to explain the primacy and motivational force of epistemic reasons, which, thereby, accounts for why beliefs change in the face of counter evidence, even unconsciously.

In the following sections I examine two objections in the literature – the first by Owens (2003) and the second by Shah (2003) – followed by my own examination of whether either of these accounts can make sense of the features of belief above.

5. Beliefs cannot aim

Owens (2003) notes that in cases of pragmatic reasoning, what Thomas Kelly (2003) calls "instrumental rationality" (Kelly, 2003:612), we are able to weigh up the

desirability of one goal against the desirability of other goals before making a decision about which goal to aim at. Thus, in working out whether to carry on working or go for a swim, my desire to work is weighed up against my desire to swim. Once I've weighed up the consequences and relative merits of both goals I then choose to pursue one or the other, or neither if some other option presents itself as an overwhelming favourite.[52]

But, says Owens (2003), with beliefs, I can't weigh the goal of Truth against my desire to feel good about myself, for example, my beliefs can have only one goal, Truth.[53] Given that beliefs can have only one goal, argues Owens, they cannot be said to aim at it.

Owens' (2003) worry is this. In the case of practical reasoning, I'm able to decide which goal to pursue. In aiming at *X* I aim *X* by choosing to pursue *X* in favour of other goals. But this isn't the case with beliefs. There is one and only one goal, Truth. What this means, says Owens is that beliefs can't really be said to have an aim. To aim at something, Owens seems to suggest, is to be able to aim at more than one thing. And since there's nothing else to aim at, because beliefs cannot have a plurality of goals, the aim of Truth is no aim at all.

Steglich-Petersen (2009) sums up this objection neatly:

> … if our adherence to epistemic norms is explained by some *aim* of ours, we should expect that we, on at least some occasions, when deciding how to adhere to the norms generated by that aim, weigh the

[52] I take it that not all aims are necessarily competing. I can have the goals of having dinner with my parents and eating fish and make them work together in order to satisfy both aims. In this instance the goals are complementary, rather than competing. However, these goals do compete with other goals, such as, say, staying at home and watching a movie.

[53] We need to be a little bit careful here. As Timothy Chan (2013) notes that "In recent years … the idea that belief aims at knowledge rather than truth has been argued for in the literature" (Chan, 2013, 10). What this means is that accounts of how beliefs explain the features of belief can postulate goals other than Truth. However, whatever this goal is, it must be sufficiently close to Truth to explain the primacy and motivational force of epistemic reasons and must aim at this goal exclusively in order to account for the motivational force of epistemic reasons. Thus, if beliefs aim at knowledge rather than Truth, then they aim at knowledge exclusively, rather than in addition to, for example, happiness. A system that aimed at knowledge and happiness wouldn't be able to explain the primacy or motivational force of epistemic reasons, or account for the formation of beliefs that, while seemingly True, nevertheless made us unhappy.

> aim of belief against other aims. But epistemic deliberation is not like that, they [Shah (2003), Kelly (2003) and Owens (2003)] say. When we deliberate about what to believe, we focus exclusively on epistemic considerations bearing on the truth of the proposition being considered for belief. (Steglich-Petersen, 2009:396)[54]

And here we hit a snag for teleological accounts, argues Owens (2003). If beliefs don't aim at anything, then all the explanatory work that's being done by the teleological relationship with Truth is undermined. What makes Aim and Purpose Accounts attractive is that they are able to explain beliefs' adherence to doxastic norms in virtue of either a constitutive aim at Truth, or as part of an inherent purpose of Truth. If Owens is correct then, *prima facie*, Aim Accounts are definitely in trouble and perhaps so too are Purpose Accounts.[55]

5.1. Owens' objection and Aim Accounts

If Owens (2003) is right and beliefs cannot be said to aim at Truth, in virtue of there being nothing else to aim at, then Aim Accounts are in serious trouble.

Aim Accounts are premised on the fact that what does the explanatory work is the aim of Truth. This constitutive aim generates certain norms, which, in virtue of the aim,

[54] As mentioned above, this direct link between truth and belief seems to be equally at play in non-deliberative contexts and so I take this objection to count similarly against instances of non-deliberative believing.

[55] In the Asbjørn Steglich-Petersen (2009) quote above, he talks about "*our* adherence to epistemic norms … [being] explained by some *aim* of *ours*" (Steglich-Petersen, 2009:396) (italics added). He goes on to say that some, like Thomas Kelly (2003), Nishi Shah (2003) and David Owens (2003), deny that we can weigh *our* aim for Truth up against other aims in cases of doxastic deliberation. The concern for me here is that Steglich-Petersen is conflating the potential aims of a believer with the aim of belief. It is patently obvious to me that as doxastic agents, even if we're engaged in doxastic deliberation, we can aim at a plurality of different targets. Imagine that I want to know if something is true (that Brazil and Spain are the only two countries to have won the Football World Cup outside of their own continents, for example). While engaged in deliberation about this project, I come across the fact that a movie I wanted to watch is about to finish screening at the cinema. It seems clear to me that in this instance, my, the believer's, aim of Truth will be weighed up against my desire to go to the movie, for example, and that I either persevere with my goal of Truth or forego it in favour of some light entertainment. In this case it seems that as a believer I can quite comfortably aim at the Truth of this claim as my goal, even if Owens' (2003) objection is correct. While beliefs may only be able to aim at one thing, Truth, this doesn't mean that I as an agent can't weigh a project of deliberation, in which the Truth of this proposition is the goal, against other goals. In these instances, where it is the believer who has the aim of Truth, we will see our aim and/goal interact with other aims and/or goals in the familiar way associated with pragmatic reasoning. Owens' objection, therefore, applies only to whether beliefs *qua* beliefs can be said to legitimately aim at Truth.

account for several features of belief. If it is impossible for beliefs to aim at Truth then Aim Accounts must be discarded.

Owens' (2003) argument seems to be this.

1. If beliefs could aim at the Truth, then they could aim only at the Truth.
2. More than one goal is required to have an aim.

Therefore,

3. Belief cannot be said to have an aim.

The first premise I accept. Given what it is to be a belief, given the constraints on what reasons we can take as reasons to believe, given the constraint on the explanations we're able to tell of our beliefs, given the involuntarism of belief and so on, if beliefs can be said to aim at anything it must be the Truth and only the Truth if we're to explain any of these features in terms of an aim of belief.

The questionable premise for me, the one that Owens' (2003) argument rests on, is the second premise: namely, that to have an aim one must be able to aim at more than one target. What this claim appears to amount to is that to be able to have an aim at all it is a constitutive condition of having an aim that there is more than one thing to aim at.[56]

Yet, despite the fact that this premise is central to Owens' (2003) argument, his defence of this premise appears to be based on an observation that in practical reasoning we see our goals interact with other goals, something that doesn't happen in epistemic reasoning.

[56] It is important to note that the claim here is not that having practical *reasons* for different goals is what makes an aim an aim, but rather that having different *goals* is the constitutive element of having an aim. Practical reasons give us reasons to aim at one goal rather than another, but are not goals in and of themselves and so are not constitutive of aims. If this is wrong, and reasons themselves are goals, then Owens' (2003) objection is easily resolved. If all that is needed for something to aim is to have competing reasons, then beliefs can aim. In the case of belief the competing reasons are epistemic reasons. While we might not be able to take on board practical reasons for believing this doesn't mean that beliefs don't face competing epistemic reasons. Given the distinction between practical and epistemic reasons I can see no reason why we should think that beliefs are required to compete with pragmatic reasons. In this case we can make sense of making an aim attribution with regards to beliefs, so long as we keep the practical/epistemic-reason distinction clear in our minds. For the objection to stand, goals must be distinct from reasons and it must be that we cannot have an aim without the constitutive requirement of multiple goals.

However, to claim that beliefs *cannot* aim because of something we observe in cases of instrumental rationality seems too quick. This correlation, in and of itself, doesn't prove that an aim requires more than one target. All it shows is that in instances of practical reasoning by human agents we have more than one potential goal.

If the constraint is conceptual, however, if there is something about being an aim that requires there to be more than one possible target, or something about goals that requires more than one of them in order for us to aim, then Owens' (2003) argument is in good shape. Anything less than a conceptual constraint, however, suggests that it would be logically possible for beliefs to aim at Truth.

In what follows I will try to spell out a conceptual argument for the second premise in Owens' (2003) argument.

Let's go back a few steps to see if we can make more sense of this objection.

I am trying to decide whether to stay in bed or to get up and do some work. In this case I have two goals for which are various conflicting reasons: sleeping in (goal) and feeling rested (reason) or getting up and doing work (goal), which will make me feel productive (reason). At this stage, I don't have an aim, unless it is to work out which of my goals to pursue.

If Owens (2003) is right, then I can only form an intention to pursue (aim at) one of these goals because there is another potential goal. What lends weight to Owens' argument is that the moment I examine any of my instrumental aims I find that there are always potentially competing goals being passed over in favour of the chosen one.[57]

[57] I have tried to come up with an example in which there are no other complementary or competing goals, without doing so to my satisfaction. Consider the following. You have been captured and are on an island where you are told you will be executed in the next 24 hours. You decide to escape. You manage to tunnel your way through the floor and now, to escape, you are forced to swim underwater to reach freedom. While you are swimming you have only one aim: to hold your breath until you reach the other side. Would this count as a candidate for someone having only one aim, thereby undermining Owens' (2003) objection? I take it that the answer is no. While you currently have this aim, you could have another. You could decide that actually the trouble isn't worth it and you're going to either stop holding your breath and let the water take you, or turn around and wait in the prison cell. I certainly

But this observation doesn't prove Owens' (2003) claim. It may just be the case that as embodied human beings we always have a plurality of goals that we may pursue; a sort of case of Jean-Paul Sartre's (1956) "slave in chains" (Sartre, 1956:703) always having more than one goal he/she can pursue. If this is the case, then it says something interesting about our experience as agents, but doesn't prove anything about aims. So, even though beliefs cannot aim at more than one thing I'm still not convinced that the acknowledgement that we as agents always have a plurality of goals is sufficient to prove that this is a necessary requirement of aims.

To show that beliefs *cannot* aim in virtue of not having more than one goal to aim at there must be something about aims or something about goals that requires there to be more than one of the latter in order to have the former.

Starting with goals, one way of understanding Owens' (2003) claim is that goals (roughly, an end result, a target, a goal or something similar) are somehow constitutive of aims. Thus, to have an aim you must have at least two goals.

If Owens (2003) is right, then in order to aim at eating an ice-cream, then it cannot be the case that this is the only thing I'm able to aim at. If I am to aim at eating an ice-cream at all, then it must the case that there are some other goals that I could aim at, even if I probably wouldn't. In other words, I can only aim at eating an ice-cream *because* I have a competing goal of, for example, wanting to lose weight.

Put like this, the claim seems exceedingly odd. What it amounts to is that aims are constituted in virtue of the fact that more than one goal exists. As mentioned earlier, while it is true (perhaps contingently) that we as agents find ourselves able to pursue more than one goal, it seems strange that having more than one goal is a necessary condition of having an aim.

The best way I can make sense of Owens' (2003) objection is as follows. To aim is to intend to reach a target. If I aim to finish my book before I go to bed, I form the

---

can't envisage a situation in which there seems to be one and only one aim. Further, to stipulate that these cases exist would seem rather question begging.

intention to complete a certain task by a certain time. The target in this case is the completion of the book by a certain time and the aim is the intention to reach this target. To aim, therefore, is to form an intention to reach a certain goal. And goals are things we can choose to aim at. Now, if goals are things we *choose* to aim at, then perhaps Owens' objection gains weight in virtue of what it means to be a choice.

We need to be careful here. The claim here is not that aims are just choices: instead the claim is that in order to aim at *x*, we must be able to choose to aim at *x*. And being able to choose *x* in any meaningful way may mean that there must be another option other than *x*. Thus, if we can't choose to aim at one goal over another, in virtue of there being no other goals to aim at, then perhaps we can't aim at all.

Let's examine this argument in a little more detail, beginning with a distinction between choices and aims. To choose something is to pick one of several options. To aim at something is to pursue a specific end. Of course there's an obvious sense in which to aim at *x* is to haven chosen *x* and to make choices in line with *x*. To aim at eating an ice-cream is to have already chosen to pursue this aim rather than the aim of losing weight, for example. But while to aim at *x* may be to choose *x* over other options there is still a clear distinction between the two.

Consider a game show where there are three different-coloured blocks, one of which has the correct answer behind it. In each turn I need to choose the block with the correct answer behind it. If I choose correctly I move on to the next round where I need to correctly guess which block has the right answer behind it and so on. If I get it right five times in a row, then I win the game.

It seems to me that choosing which block to pick is a random selection that in and of itself aims at nothing. While I aim at winning the money, or aim at getting this guess right, this intention drives my choices. If my intentions were different, my choices would be different. So, while my choices are based on my aim to win, they are nevertheless distinct from my aims.

This is all slightly unwieldy. Perhaps the simplest way to understand the distinction is this. If beliefs can aim at Truth, in way that explains the motivation of epistemic

norms etc., it's not that they choose Truth instead of something else. Instead by aiming at Truth they engage in certain things to try to meet that target. It is this aiming rather than the choice that generates the norms that do so much explanatory work in Aim Accounts. A choice, in and of itself, it seems would produce no such norms unless it is coupled with an aim to follow through with that choice. To have an aim, however, means to have already made a choice about which end to pursue, coupled with an intention to pursue that end.

If to have an aim requires being able to make a choice about which course of action to pursue, and all other options are removed, then it seems that I cannot make a choice at all. And if I cannot make a choice about which end to aim at, then perhaps I cannot be said to aim at all.

Peter Van Inwagen (2009) notes in his chapter on free will and determinism, that if determinism is correct, then free will, the capacity to make effective choices, is an illusion (Van Inwagen, 2009:269).

> If Jane is trying to decide whether to tell all or to continue her life of deception, she is in a situation strongly analogous to that of someone who is hesitating between forks in a road … To say that one has free will is to say that when one decides among forks in the road of time (or, more prosaically, when one decides what to do), one is at least sometimes able to take more than one of the forks. Thus, Jane, who is deciding between a fork that leads to telling all and a fork that leads to a life of continued deception, has free will (on this particular occasion) if she is able to tell all and is also able to continue living a life of deception. One has free will if sometimes more than one of the forks in the road of time is "open" to one. One lacks free will if, on every occasion on which one must make a decision, only one of the forks before one—of course it will be the fork one in fact takes—is open to one. If John is locked in a room and doesn't know the door is locked, and if he is in the process of deliberating about whether to leave, one of the alternative futures he is contemplating—leaving—is, in point of fact, not open to him, and he thus lacks free will in the matter of staying or leaving. (Van Inwagen, 2009:253-254)

In the case of determinism, the ability to make effective choices is nothing more than an illusion. There is only one thing that an agent can do, and so to say that he/she in any real sense "chose" to do *x* is just fantasy. If aims require choices, then beliefs cannot aim, because beliefs can't choose to aim at Truth in favour of something else.

The argument then is this. If beliefs choose Truth as an end to pursue, then it would seem that on at least some occasions beliefs would choose something else. But this is impossible. An account that sees Truth as one goal among a number of other instrumental goals that beliefs could pursue could not account for the features of beliefs. Thus, if beliefs do aim at Truth, then that would be the only thing they could aim at. But if aims require the ability to choose between different aims, then beliefs cannot aim. If beliefs cannot aim, then Aim Accounts must be abandoned in the face of Owens' (2003) objection.

For what it's worth, I am not at all convinced that this cashing out of Owens' (2003) objection is right; for two reasons. One, I'm not certain that what I've laid out above is what Owens has in mind. Perhaps there is something else going on in Owens' argument that clearly explains why aims would need a plurality of targets in order to exist. Two, I'm not sure that a lack of choices necessarily undermines the capacity to aim. While it seems right to think that we need to choose between options in order to aim at one option rather than the other, I worry that I'm merely repeating something like the correlative claim that I suggested Owens might be making. Is it a necessary feature of aims that there is necessarily an antecedent choice? I don't know.[58]

I feel that more work needs to be done to show that aims require a plurality of goals in order to exist. Yes, we see our practical reasoning work within the framework of

[58] There is a further concern to Owens' (2003) objection that I'm ignoring here. Gregory Kavka (1983) argues that intentions aren't things that we can form at will. If we can't choose to form intentions and aiming at something is predicated on having an intention to aim at something, then perhaps we don't need to be able to choose between goals to have an aim. However, I'm not sure where I stand on Kavka's argument or what ramifications an involuntarism of intention (if any) has on the concept of aims. If an intention precedes a choice, then perhaps Owens' objection falters (assuming my story about choices is Owens' claim). However, if a choice precedes an intention, then it seems as though, while potentially complicating matters, Owens' objection would still stand.

multiple goals, but I'm not sure that this means that we need to have this plurality in order to be able to aim at one of them.

However, this is as far as I will take this objection. What I wanted to show is that if Owens (2003) is right, then Aim Accounts are undermined as an account of the truth-belief relationship.

It is at this point that my distinction between Aim Accounts and Purpose Accounts becomes important. While Owens' (2003) objection may well undermine Aim Accounts of the truth-belief relationship, as I argue below it cannot undermine Purpose Accounts; highlighting the importance of making the distinction.

5.2. Owens' objection and Purpose Accounts

Owens (2003) argues that since there is no plurality of goals for beliefs, talk of aiming is misguided. As argued above, while I feel that the argument requires greater unpacking, there does seem to be something here. If we need to be able to choose between two different options in order to aim at one rather than the other, then being able to aim requires being able to choose between two options. Since beliefs cannot choose between Truth and something else, beliefs cannot choose and, so, cannot aim.

However, while an Aim Account would require the ability to aim in the sense laid out above, there is no reason to think that something that has a purpose needs to have a plurality of purposes in order to have even one. In fact, there is no reason to think that something that has a purpose needs to aim at all.

To have a purpose is for something to have an inherent goal. In toasters that goal is to cook bread, in kettles it is to boil water, and so on.

In virtue of having a purpose certain things need to happen for that purpose to be discharged. In the case of a kettle, for example, water must be heated until it reaches boiling temperature. Clearly in this case we can say of the heating element that it aims at heating the water until it reaches boiling point. But this would be to use metaphorical language for ease of use rather than because heating elements

intentionally aim at anything. Heating elements, to steal Wedgwood's (2002) line, "are not little archers armed with little bows and arrows: they do not literally 'aim' at anything" (Wedgwood, 2002:268).

The distinction lies in the difference between aiming at something and something having a purpose. In the case of the latter, the processes that need to happen in order for that purpose to be fulfilled occur not as a result of an intentional aim to reach a target, but as part of the workings of that purpose. Thus, while we can talk about the heating element aiming at heating water, this aiming is a norm (for the heating element must or ought to heat the water to boiling point) generated out of the purpose of the kettle.[59]

The question is whether in the case of purposes whether the processes that need to occur in order to fulfil that purpose need to "aim" at anything else? The answer seems to be quite obviously, no. When something has a purpose then the norms generated out of that purpose, the processes and things that need to occur in order to fulfil that purpose, are as a result of that goal only. What this shows is that things that "aim" at fulfilling the purpose are just the norms that are generated out of that purpose and aren't aims in the sense detailed in the discussion of Aim Accounts above. Because these processes aren't aims in this sense they don't need to choose between options. They are the concomitant side effects, if you will, of the purpose of the system in question. Further, it doesn't seem that the purpose needs be able to engage with other purposes in order to be a purpose. That a kettle has the purpose of boiling water doesn't mean that it would have had to have had other possibilities in order to exist.

"Ah," an imaginary objector might say "but even though the kettle may have no other purposes, the person who made that kettle surely did. Doesn't this mean that we need to have more than one goal in order to have a purpose?"

But, while the person who put the kettle together could have chosen to do something else, or put the kettle together in some other way, the actual system has only one purpose, with everything that occurs to fulfil the purpose as a result of that purpose.

[59] A potentially interesting side point is that in Aim Accounts the aiming generates the norms. In a Purpose Account the things that could be said to aim *are* the norms.

In this case, in the case of purposes, we can speak quite comfortably about the aim (understood as part of the purpose) of belief, without needing to worry about there being more than one goal for beliefs to aim at.

Arthur Collins (1978), in his discussion on teleology, mentions the human homeostatic system (Collins, 1978:543). This bodily system, which controls our body temperature, is seemingly both teleological and single-minded in its purpose. The sole goal of the system is to ensure that our bodily temperature falls within a certain range. And the compensatory behaviour (sweating in excessive heat, shivering with cold, etc.) is explained in virtue of this goal. In this instance, we have an example of a system that regulates in virtue of having a purpose and yet has no competing or complementary aims.

Similarly, if Truth is the purpose of belief, then one of the advantages of this type of account is that it is able to avoid the worry that Owens (2003) has above about needing a multiplicity of goals in order to make sense. If this is right, then Owens' objection, while potentially fatal to Aim Accounts, fails to undermine Purpose Accounts.

6. The Teleologist's Dilemma

The problem with teleological accounts, says Shah (2003), is not whether they can aim, not that they aim at Truth, but that it's not possible to aim at Truth strongly enough to capture transparency, but weakly enough to capture motivated beliefs (Shah, 2003:460-465). Let's briefly go over these two features of belief again – starting with transparency – before considering Shah's objection.

I argued in Chapter 2 that there are three possible ways of understanding what Shah (2003) means by transparency: one, transparency is an acknowledgement of the direct relationship between belief and truth; two, transparency is the acknowledgement of the primacy of epistemic reasons for belief; or, three, transparency is the claim that there is a direct link between epistemic reasons and beliefs such that the former have a direct impact on the latter. While each of these three features of belief is distinct, I believe that the argument below would apply to all three. In this section, therefore, I use Shah's "transparency" to cover all three.

Motivated believing is when an agent forms or maintains a belief for non-epistemic reasons – traditionally motivated by a desire or fear. What is important to remember about motivated beliefs is that the agent who has them is convinced of their Truth while he/she has them and should, if the belief maintains, tell an epistemically rationalising story about them. Thus, while motivated beliefs are caused by non-epistemic reasons, from the agent's perspective these reasons are opaque to the believer and the motivated beliefs are as legitimate as any other belief that he/she has. Should he/she come to see that his/her motivated belief is based on a non-epistemic reason that belief will disappear.

Shah (2003) recognises that accounts of the truth-belief relationship need to make sense of both features of beliefs. The problem, Shah argues, is that teleological accounts are "caught on the horns of a dilemma" (Shah, 2003:461): either the link with Truth must be strong enough to explain transparency, in which case it eliminates the possibility of motivated believing, or the link must be weak enough to account for motivated believing, in which case it rules out transparency (Shah, 2003:460-465).

> On one horn, the teleologist must allow the disposition that constitutes aiming at truth to be so weak as to allow paradigm cases in which beliefs are caused by such non-evidential processes as wishful-thinking, in which case he cannot capture the exclusive role of evidence in one particular type of belief-forming process, reasoning. On the other horn, in order to account for the exclusive role of evidence in reasoning about what to believe, the teleologist must strengthen the disposition that constitutes aiming at truth so that it excludes the influence of non-truth-regarding considerations from such reasoning. However, by strengthening the truth-aimed disposition, the teleologist cannot accommodate the cases of wishful thinking, in which non-evidential factors clearly exercise influence over belief. (Shah, 2003:461)

In this section I lay out a version of Shah's (2003) objection. Shah argues that transparency is a feature only of deliberation. Shah, therefore, thinks that teleological accounts are only undermined by his dilemma when it comes to their inability to explain a feature of deliberation; that in other types of believing we can still use a teleological account of the truth-belief relationship. As he puts it: a teleological account "seems to explain why evidence-driven processes such as perception normally induce belief" (Shah, 2003:473).

However, as I have argued in Chapter 2, transparency is a feature of all beliefs. Thus, if teleological accounts are unable to make sense of transparency, then this undermines their ability as accounts of the truth-belief relationship, *simpliciter*. With this in mind, here is my version of Shah's (2003) objection.

Shah (2003) argues that when it comes to believing, particularly in the case of deliberative believing when one is focussed on believing correctly, transparency is highlighted as a feature of belief and is governed by the relationship that a belief has with truth.

That belief has these truth-tracking features built into the very fabric of what it means to be a belief suggests a very strong link with Truth. A link that is so strong that

questions about whether to believe that *p* are answered only by reasons that answer the question whether *p* is true. The former question is transparent to the latter one.

But if there is such a strong link with truth such that the *only* reasons we can take on board for beliefs are epistemic reasons, how do we ever get to the point where we form motivated beliefs? Motivated beliefs, as noted above, are characterised by the fact that they are beliefs that are formed by, and/or sustained by, non-epistemic reasons. If the link between belief and truth is so strong as to account for transparency, it seems as though motivated beliefs couldn't be formed at all.

> If the disposition constitutive of belief were so strong as to rule out the influence of any other factors, then any cognition that came to be influenced by non-evidential factors would cease to be a belief. The wishful thinking cuckolded husband who believes his wife to be faithful or the racist whose dislike of African-Americans leads him to fall prey to poor sociobiological or other "scientific" arguments wouldn't be irrational, but impossible! (Shah, 2003:465)

Similarly, Shah (2003) argues, if we allow a link weak enough to account for motivated beliefs, then we can't make sense of the fact that only what we take as epistemic reasons count as reasons for belief (Shah, 2003:460-465).

To reiterate an important, but potentially tired, point, we're not interested in a link between believers and truth. As I've argued for previously a believer can either aim at Truth or at something else (as is the case of Pascal's Wager). Further, as argued for in Chapter 2, the norms that are generated out of believer's goals are irrelevant to the truth-belief relationship. The link we're interested in is the link between belief and truth. So, if Shah (2003) is right and teleological accounts are on the horns of a dilemma, then it is because of the inherent relationship between beliefs and truth.

What highlights transparency as a claim about dispositional as well as occurrent beliefs is that even when we form beliefs unconsciously these beliefs follow truth-tracking norms such that they regulate for truth. Thus, my beliefs unconsciously adjust, change, maintain in the light of epistemic reasons and do so as a result of

transparency without this needing to be a case of conscious deliberation. That beliefs can change unconsciously, taking only what appear to be epistemic reasons into consideration – and do so directly – highlights what's at stake: an account of belief that makes sense of this strong concept of Truth in order to account for transparency. But, given transparency, how do we make sense of motivated beliefs? The challenge for teleological accounts, therefore, is to explain how the link between belief and truth is both strong enough to account for transparency, but weak enough to account for motivated beliefs.

Below I examine both Aim Accounts' and Purpose Accounts' ability to defend their version of the teleological account in the face of this objection. I argue that while this challenge is a potentially fatal objection to Purpose Accounts, Aim Accounts are able to deal with this challenge relatively easily.

6.1. The Teleologist's Dilemma and Aim Accounts

Cashing out Shah's (2003) objection to teleological accounts in terms of aims, the concern for Aim Accounts is that the target is either too strong or too weak to account for certain features of belief.

I confess that if we're talking about Aim Accounts, I'm not entirely sure how to make sense of this objection. A target is a target and can seemingly neither be too strong or too weak a thing. While an agent, or sub-intentional agency, may have a strong or weak ability to *hit* the target it seems nonsensical to talk about having too weak or too strong a target. So how do we cash out this objection in the case of Aim Accounts?

Perhaps the way to do it is to talk about the conception of truth that beliefs aim at. As mentioned above, it seems as though beliefs can aim either at truth or at Truth. I argued above that the target for Aim Accounts must be Truth, rather than truth. If beliefs aim at truth, then we could not account for beliefs regulating for truth as the target would already be what the belief had reached. As a result, the belief would have no need to try to hit anything else. In this case, in the case in which beliefs aim at truth, epistemic reasons wouldn't have any motivational force on our beliefs.

Now, while regulating for truth in the face of counter evidence is not strictly transparency I think it is close enough to be part of the version of this objection. Manipulating Shah's (2003) objection slightly, the objection would be that the conception of truth, in an Aim Account that postulates truth as the target of belief, is too weak to account for transparency-like features of belief. And since it cannot account for these features of belief, Aim Accounts that aim at truth should be abandoned.

The solution to this problem, I argued above, was to postulate that Aim Accounts must see beliefs as aiming at Truth. However, the problem here may be that if beliefs aim at too strong a conception of Truth, then perhaps they can't account for motivated beliefs. If this is the case, then Aim Accounts are squarely on the horns of Shah's (2003) dilemma and ought to be abandoned as a result.

If beliefs aim at truth, then I think that Shah's (2003) modified objection has Aim Accounts on at least one horn of the dilemma, as we couldn't then account for transparency or transparency-like features of belief. Aiming at truth provides no capacity for beliefs to take epistemic reasons on board and so the motivational efficacy of epistemic reasons would be completely undermined.[60]

However, while Aim Accounts that aim at truth are potentially in trouble, I think Aim Accounts that see beliefs aiming at Truth are able to account for both doxastic phenomena.

Let's begin with transparency and transparency-like features of belief. As I argued above, a constitutive aim of Truth would require beliefs to take only epistemic reasons as reasons for belief. If I'm aiming at Truth, then the only way to do so would be to ensure that the reasons I take on board are reasons that are Truth-directed. Similarly, there would, in terms of the constitutive aim, be no coherent framework in which I weighed epistemic reasons for *p* against instrumental reasons for *p*. Given the

[60] For what it's worth, I'm not entirely sure that an Aim Account that aims at truth could account for motivated beliefs either. What I am fairly certain of is that an Aim Account that postulates truth as the aim of belief cannot account for transparency, or transparency-like features of belief.

constitutive aim of belief, the only way to achieve this aim would be to take epistemic reasons and only epistemic reasons as reasons for *p*.

These reasons, as a result of necessarily excluding other reasons along with the aim to hit the Truth, would mean that epistemic reasons have a direct impact on what we believe, and, thus transparency is explained.

However, to aim at Truth does not mean to always succeed. I can aim and miss, and if I'm not really paying attention still think I've successfully hit the target. Similarly, it may be the case that I'm blinded by rage or desire and as a result shoot and miss, but still think I've hit. This seems to describe a classic case of motivated believing. Wishful thinking, in which one believes something not because one has good epistemic reason to, but because one wants to, seems to be a clear case in which the archer of belief, blinded by hate or passion misses the target of Truth, but "thinks" the correct target has been hit.

Williams (2002) suggests that:

> The phenomenon of wishful thinking is very basic and not a great mystery: the steps from its being pleasant to think of P, to its being pleasant to think that P, to that thinking that P, cover no great psychological distance. (Williams, 2002:83)

In this case, while a belief aimed at Truth it missed resulting in a motivated belief. This is not to say that the belief failed to aim at Truth, but simply that, as a result of other internal and external factors, it failed to reach the correct mark. Thus, while Aim Accounts are potentially in trouble from Owens' (2003) objection above, I think they are safe from Shah's (2003) dilemma. They are able to account for transparency and transparency-like features of belief in virtue of a constitutive aim and make sense of motivated beliefs in virtue of having inadvertently missed the target.

However, while Aim Accounts are I feel exempt from this challenge, Purpose Accounts fare far worse in the face of this challenge.

6.2. The Teleologist's Dilemma and Purpose Accounts

I argued above that if beliefs have a purpose in terms of Williams' (1973) aim-at-truth metaphor, then it must be Truth, rather than truth, otherwise Purpose Accounts couldn't account for the fact that beliefs regulate for truth. However, as I shall argue below, neither conception of truth would save Purpose Accounts from Shah's (2003) dilemma. I begin by examining whether a belief that has truth as its purpose can escape Shah's dilemma, before examining whether a belief that has Truth as its purpose can escape this objection. I then consider whether there might be a middle road of truth, stronger than truth, but not as epistemically strict as Truth (truth+), that could account for both transparency and motivated beliefs.

Let's start by examining whether Purpose Accounts that have Truth as their purpose can avoid Shah's (2003) objection by going back to the trap of belief. Consider in this case that the trap has been specially designed to catch unicorns. This trap, to ensure that it doesn't catch anything that's roughly the right size and shape of a unicorn, is specially calibrated to take certain things into account in order to ensure that it gets unicorns, rather than, say, donkeys.

To catch a unicorn, therefore, the trap must take only certain types of evidence into consideration: namely, evidence that pertains only and directly to whether the animal in the trap is actually a unicorn. In this case, the question 'whether to catch *p*' is transparent to the question 'whether *p* is a unicorn' and only reasons that answer 'whether *p* is a unicorn' are relevant to the question 'whether to catch *p*'.

In this analogy, a unicorn is Truth and the belief is the specially designed trap that springs only once Truth has been reached. Further, in order to catch the Truth, there are only certain reasons that the belief can take on board in virtue of the purpose of belief. And these reasons, in the case of belief, must be epistemic reasons. This purpose of Truth explains why the deliberative question 'whether to believe that *p*' is transparent to the question 'whether *p*' in much the same way that question 'whether to catch *p*' is transparent to the question 'whether *p* is a unicorn': only reasons that answer the latter are relevant to the former.

Thus, if beliefs have the purpose of Truth, then we can quite easily understand transparency. The transparency comes about as a result of the purpose of belief.

Now, let's consider beliefs that have truth as their purpose. In terms of the analogy above, this trap wouldn't catch only unicorns it would catch what it took to be unicorns. In much the same way that a belief that has truth as its purpose doesn't have actual Truth as its purpose but only what an agent takes to be True (the truth).

Could this trap make sense of transparency, or transparency-like features of belief? No.

Catching non-unicorns would mean that the trap would admit to reasons that don't directly answer the question 'whether *p* is a unicorn'. If my trap can spring when a donkey is in it, for example, then it can only do so in virtue of the fact that it takes reasons other than reasons that answer whether the animal is a unicorn into account. If this is the case, then this trap and this model of belief cannot account for transparency.

What if Purpose Accounts have Truth as their goal? If the end goal is Truth, then only epistemic reasons for *p* could count as reasons to believe that *p*. Further, given the inherent purpose of belief, these epistemic reasons would have a direct link with belief such that in seeing that *x* was a sufficient reason for the Truth of *p* you would believe that *p*. For beliefs to take on board reasons other than epistemic reasons would be to fail to meet the purpose of Truth in much the same way that a unicorn-trap that captures donkeys is failing the purpose of being a unicorn-trap. If this is right, then it seems to me that a Purpose Account that postulates Truth as the purpose of beliefs would have no problems accounting for transparency.

But if Purpose Accounts have Truth as their goal, then it's hard to see how we could ever be wrong. If our unicorn trap is designed to capture only unicorns, taking only unicorn-directed reasons as reasons for whether to close, how could it admit to other, non-unicorn-directed, reasons? Put differently, if we have a unicorn trap whose only purpose is capturing unicorns and it is so calibrated as to make sure it captures unicorns, how could we ever catch a donkey? By analogy, if our beliefs are only formed when a candidate is True, thereby admitting only to epistemic reasons, then

how can we make sense of a belief ever taking on board non-epistemic reasons? And how can we make sense of ever having motivated beliefs?

One answer would be that when an agent takes something as True, he/she satisfies the purpose of belief such that the trap springs accidentally. In this case, while the belief has achieved the purpose of catching only the Truth, it has done so in virtue of the fact that the believer has mistaken some non-epistemic reason as an epistemic reason.

However, this answer can't work. The question that Purpose Accounts would need to answer is how the believer mistakes some non-epistemic reason for an epistemic reason. As argued previously, the purpose of a belief is not same as the purpose of the believer and a believer doesn't have direct control over his/her belief. A Purpose Account suggests that there is a relationship between belief and Truth, such that beliefs have Truth as their purpose. In the case in which an agent takes a non-epistemic reason as an epistemic reason this would need to be explained in terms of a relationship between beliefs and truth where beliefs have Truth as their goal. If it is beliefs themselves that have the goal of Truth, then how do we make sense of them admitting non-epistemic reasons? We can't. To mix analogies somewhat, the trap is designed to take only a certain kind of reason into account in order to ensure that when it snaps shut it doesn't capture a donkey.

It seems as though Purpose Accounts are squarely on the horns of Shah's (2003) dilemma. A weak account of truth, where the purpose of belief is what an agent takes to be True, allows for motivated beliefs, but is unable to account for transparency. A strong account of truth, where the purpose of belief is the Truth, allows for transparency, but cannot make sense of motivated beliefs. What Purpose Accounts need is a purpose somewhere between truth and Truth; somewhere between what an agent takes as Truth and the actual Truth, a sort of truth+.

What would truth+ look like? Let's go back to the trap. This trap would ensure that it didn't catch things that were obviously not unicorns (cats, for example), but would admit to the fact that the trap wasn't perfect. That in some instances a horse, or even on occasion a donkey, could get trapped. Truth+ then would be a conception of truth

that would be just weak enough to account for motivated beliefs (donkeys) and perhaps just strong enough for transparency.

But, this doesn't work. Any account that allows for non-epistemic reasons can't, in and of itself, explain transparency at all. A Purpose Account that has truth+ as the goal of belief would still admit to the fact that we get it wrong and get it wrong because our beliefs are able to take non-epistemic reasons into consideration. But this is precisely what transparency denies.

Shah's (2003) objection, therefore, while not a problem for Aim Accounts, poses a real threat to Purpose Accounts. Purpose Accounts that postulate truth or truth+ as the inherent purpose of belief can't explain transparency. Purpose Accounts that postulate Truth as the purpose of beliefs can't make sense of motivated beliefs.

Before moving on to some objections of my own, what these two objections highlight is the importance of the distinction between the two types of teleological account. As we've seen above, each of the objections appear to be germane to only one of these accounts rather than the other.

What this means is that those interested in defending Aim Accounts, for example, would need to respond to Owens' (2003) objection and needn't respond to Shah's (2003) objection. Purpose Accounts, on the other hand, could avoid getting entangled by Owens' objection and focus on how to respond to Shah's dilemma. The failure to appreciate this distinction has made things far murkier than they need to be when talking about teleological notions of the aim-at-truth metaphor.

Below I examine other objections to the two teleological accounts. In Section 7 I consider other potential problems with Aim Accounts, before moving onto other potential problems with Purpose Accounts in Section 8.

7. Aim Accounts and other problems

If Owens (2003) is correct then there is a deep conceptual problem with Aim Accounts: namely that in order to aim at all the aimer needs to be able to aim at more

than one target. If this is right, Aim Accounts are non-starters and we need to think of other ways of understanding the truth-belief relationship. However, if there is a way around this objection, then perhaps Aim Accounts are the way to cash out the aim-at-truth metaphor.

In this section I want to suggest that even if we can get round Owens' (2003) objection there are other problems with Aim Accounts that undermine them as the correct account of the truth-belief relationship.

The first is that if Aim Accounts are correct, then it's hard to understand how *idées fixe* could ever exist. The second is that Aim Accounts lose what it means to be a belief. Aim Accounts, in virtue of being Aim Accounts, place belief and Truth in tension with one another such that beliefs aim at Truth. This tension explains why, for example, beliefs tend to regulate for truth. However, in creating the explanatory tension we cannot account for another, crucial, feature of belief: namely the direct link between beliefs and truth. The inability to explain this feature means that Aim Accounts lose what it is to be a belief.

7.1. Aim Accounts and *idées fixe*

In Chapter 2 I argued that there are at least four instances of *idée fixe*: bigoted beliefs; delusions; unanswerable and yet unbelievable argument; and the cases in which I was going to believe it anyway.

The claim in this section is that if beliefs themselves aim at Truth we should not encounter evidence-resistant beliefs. The reason is simple. Our very beliefs aim at Truth and in the face of continuous and epistemically valid counter evidence they should change in response to this evidence in order to be said to aim at Truth.

I argued above that Aim Accounts are able to make sense of getting it wrong. In this case while the belief aims at Truth it has missed its target. What supports this explanation of error is that when we see that our belief is wrong our beliefs tend to regulate in the face of the counter evidence; which is explainable in virtue of the constitutive aim of belief.

However, *idées fixe* are characterised precisely by a continued resistance to counter evidence. If beliefs aim at Truth, then, while we can make sense of an occasional and temporary miss, how is it possible that these beliefs fail to fulfil their constitutive aim over a period of time in which the belief is constantly subjected to counter evidence?

I consider two possible answers proponents of Aim Accounts could give. However, I believe that while these might do some explanatory work some of the time, these solutions ultimately fail to save Aim Accounts from this objection.

The two potential ways of understanding *idées fixe* in terms of an Aim Account are:

1. As in the cases of doxastic error, other commitments compromise a belief's inherent drive towards Truth.
2. The presence of an *idée fixe* indicates that the belief is damaged or broken such that it can no longer effectively aim at Truth.

In what follows I examine point 1's ability to explain each of the *idées fixe* in terms of an Aim Account before considering 2's ability to do the same. I argue that point 1 fails to make sense of any of these *idées fixe*, while point 2 might only be able to explain believing in the face of an unanswerable yet believable argument.

Let's look at the situation described by point 1. A strong enough fear of the consequences of responding to Truth considerations or the desire to maintain some sort of mental *status quo* can see us doing all sorts of weird things.

One of the features of what are sometimes called self-limiting beliefs (unjustified negative beliefs about ourselves) is that agents who have no evidence to justify these self-limiting beliefs go out and create evidence to justify their beliefs. In these instances, the agent's goal of believing something (false) about him/herself appears to compromise a belief's ability to aim at Truth.

In this case, looking at this phenomenon through the lens of an Aim Account, the explanation is that the agent creates something that resembles a "Truth", which throws

off a belief's inherent aim of Truth. The belief's ability to appreciate the counter evidence is compromised by the conflicting goals of the *believer*.[61]

Similarly, Aim Accounts could claim, one can imagine victims of *idées fixe* focussing on only certain evidence in order to create, in a very twisted way, the required Truth(s) needed for beliefs to aim at. In this case, while the belief has failed to reach the target of Truth, this has happened as a result of an agent's other commitments obscuring the target.

Let's go back to Kwame Appiah's (1990) racist, Professor Shockley, who continues to believe in the inferiority of other non-white races in the face of overwhelming counter evidence.

How do we make sense of this *idée fixe*? Given Shockley's need to maintain his belief he gives credence to evidence that supports his claim, while finding ways of rationalising away the counter evidence.

Further, I take it that beliefs can only respond to evidence that they actually encounter. Thus, if I don't ever come across evidence of the Truth of *p*, irrespective of the Truth of *p* and the inherent aim of belief, I will never come to believe that *p*. Thus, what has happened in the case of an *idée fixe* is that the believer unconsciously manipulates what he/she has access to. In doing so, the agent creates a situation in which the belief cannot gain access to the Truth. This compromises the ability of the belief to fulfil its inherent aim. So, while beliefs still aim at the Truth, they are compromised in their ability to hit the target.

If this route of saving an Aim Account from the problem of *idées fixe* is right, then the converse must be true. Namely, if Shockley's evidence-clouding commitments fell away we should see his belief respond to the evidence.

[61] This type of move understandably brings with it concerns about the nature of self-deception: how it is possible; how it is that an agent can deceive him/herself and not be aware of it; and so on. While these are interesting issues and would need to be explored in detail by anyone adopting this type of strategy to defend the Aim Account from concerns raised by *idées fixe*, my intention here is simply to highlight ways out of a potentially difficult situation for proponents of an Aim Account.

Encouragingly for Aim Accounts we do see this happening. Occasionally *idées fixe* do disappear. An agent comes to a realisation or comes across evidence that he/she cannot rationalise away and so changes his/her belief. An Aim Account in this instance seems well-placed to explain this behaviour.

That said, there are still problems here for Aim Accounts. If beliefs have a constitutive aim of Truth, then in the face of the evidence we should expect to see these bigoted beliefs behave like so many of our other beliefs and regulate for truth automatically. Now, while it may be the case that we can avoid some evidence some of the time, or in some limited instances, all of the evidence all of the time, it seems a stretch that everyone with a bigoted belief of the *idée fixe* variety fails to encounter any of the counter evidence all of the time.

Further, in the cases in which the bigot fails to avoid all of the evidence all of the time we don't, in my experience, see the relevant beliefs regulate for truth. The belief persists *despite* the fact that the agent has seen the counter evidence.

Given an Aim Account, coming across clearly Truth-directed, overwhelming evidence to the contrary means expecting that belief to change irrespective of the other potentially evidence-clouding commitments an agent has. While we might on occasion be able to point to evidence-clouding commitments that trick belief's archer into getting it wrong, I can't see how this could always be the case. So when these beliefs don't change in the face of overwhelming counter evidence to the contrary, this suggests that an Aim Account with the target of Truth cannot be right.

The same is true in the case of delusions. Given a constitutive aim of Truth how do we make sense of delusions? Appealing to other mental states that cloud the issue may explain simple error, but cannot, it appears, explain away beliefs that continue to be held in the face of consistent and overwhelming counter evidence. In these instances, the constitutive aim of belief should operate in such a way that these states could not be possible.

In addition to being unable to account for those *idées fixe* where there are other (instrumental and psychological) reasons that could potentially overwhelm a belief's

inherent aim of Truth, Aim Accounts cannot explain those *idées fixe* where there appear to be no other non-epistemic commitments to the evidence-resistant belief.

When one has a belief that one was going to have anyway, this appears to be a belief where the truth of a propositional claim is held onto without other potentially doxastic-destabilising fears, desires and psychological influences being involved. In this case, it is hard to understand how an Aim Account can make sense of these at all. Assuming that these beliefs are not motivated beliefs that are maintained in the face of counter evidence, then how do we make sense of them maintaining, given an Aim Account? We can't. Even if there are non-epistemic reasons that get muddled up in these, the ability for Aim Accounts to handle these, as shown above, is unsatisfying.

Finally, even if somehow we can show that the non-epistemic reasons are sufficient to completely impinge on the constitutive and inherent aim of Truth, the *idées fixe* where one finds an argument unanswerable and yet unbelievable renders, I think, a decisive blow to Aim Accounts.

In these cases the epistemic reasons against the agent's belief are being appreciated by the agent. The agent might even think that the *idée fixe* should shift and yet, despite this, the belief maintains. Again, while our belief may be maintained by non-epistemic reasons – as is Ribeiro's (2002) contention – a belief that aims at Truth should see, at the very least, a suspension of belief.

In these cases there is no conflict between the aim of belief and the instrumental goal of the agent (namely to believe correctly) and yet the belief remains. A belief that aims at Truth should, in the face of an unanswerable argument, at the very least suspend itself in virtue of the constitutive aim of *that* belief. That this doesn't happen suggests that Aim Accounts are poorly placed to explain this feature of belief.

As it stands, appealing to non-epistemic factors as an explanation for the *idée fixe* doesn't work. While we might be able to explain some relative resistance some of the time in some cases, it is too far a stretch to think that in all cases of *idées fixe* this strategy works. Cases where an agent finds an argument unanswerable and yet

unbelievable, specifically, are cases in which, given a constitutive aim of Truth, we should see our beliefs suspended in the face of the counter evidence.

Perhaps the route to go is to claim of *idées fixe*s that they are broken beliefs: that somehow the Truth-aiming mechanism has broken down, causing us to have beliefs that maintain in the face of counter evidence. However, I don't think this strategy works.

Let's consider how this move might proceed. Imagine someone who believes that Rio de Janeiro is the capital of Brazil. Despite being shown that absolutely every single reputable source lists Brasilia as the capital, he/she continues to believe that Rio is the capital. It seems in these cases we might simply accept the fact that this person is rationally broken. The counter evidence is incontrovertible. The agent has been shown this counter evidence and seems to understand the challenge it places on his/her belief and yet, despite this, fails to be moved by the counter evidence.

The situation described by Erik Nisbet and Kelly Garrett (2010), cited in Chapter 2, where people believe that Obama wasn't born the US, despite conclusive counter evidence to the contrary, could be said to be in this position.

Given the weight of the counter evidence the continued reluctance of the belief to change might best be explained by the suggestion that the belief's aiming mechanism is broken. If this is the case, then *idées fixe* are explained in terms of the fact that the belief in question is broken.

Starting with bigoted beliefs there are two reasons why we might want to avoid claiming that the beliefs are recalcitrant because of a broken aiming mechanism. The first is that if we accept this move it seems we cannot hold the racist, sexist, etc., responsible for his/her belief. If the racist continues to hold his/her racist belief in the face of counter evidence and the explanation for this *idée fixe* is a broken aiming mechanism, it would be harsh to hold him/her responsible for his/her continued belief. While this doesn't mean that this explanation for why beliefs persevere is wrong, it is a potentially worrying consequence of this view. The second is that bigoted *idées fixe* behave in all other ways, other than their recalcitrance, like every other belief. The

bigot when questioned about his/her reasons for his/her belief will, assuming they are being truthful and so on, provide an epistemically rationalising story for his/her belief. The bigot will see his/her belief as being True not for any extrinsic reason that's transparent to him/her but because he/she takes the proposition as True. And, perhaps most importantly, should the agent come to see that his/her belief is based on or maintained by non-epistemic reasons, he/she will give up that belief.

If the reason for the *idée fixe* is that it cannot aim at Truth, then how do we explain the fact that the belief adheres to the other doxastic norms? Following other doxastic norms suggest that in every other way the belief-aiming mechanism is in fine shape. Appealing to a broken aiming mechanism to explain this failure to regulate for truth, therefore, looks very *ad hoc*.

While an Aim Account is able to make sense of every other one of these features, it fails to make sense of the recalcitrance of this type of *idée fixe*. Postulating that the mechanism is broken so as to account for this recalcitrance might work if the belief stopped following all of the doxastic norms, but it doesn't. What this suggests is that it's not the mechanism that is broken, but rather the account of the truth-belief relationship.

In the case of delusional beliefs the claim that the aiming mechanism of the belief is broken seems apt. In these cases the beliefs don't appear to correspond with reality at all and these beliefs, therefore, seem like prime candidates to be described as having a broken aiming mechanism. However, despite the potential attraction here what highlights the delusional belief as a belief is that other than the complete refusal in *most* instances to cohere with what we think is reality, these beliefs can behave in other ways that are consistent with what it means to be a belief.

Patients suffering from delusions do tell epistemically rationalising stories about their beliefs. While those epistemically rationalising stories may seem completely out of touch with reality they nevertheless follow the doxastic norm. The agent with the delusional belief thinks the believed proposition is True; again, not for any extrinsic reason that's transparent to the agent, but because he/she takes it as True. We can

make sense of the other features of belief in virtue of a constitutive aim of Truth. But we cannot account for the recalcitrant nature of the belief.

As in the case of Aim Accounts trying to explain instances of bigoted beliefs it seems that proponents of Aim Accounts are caught on the horns of their own dilemma. They must either give up being able to explain why *idées fixe* adhere to other doxastic norms, or they must give up being able to explain *idées fixe*. Accepting either horn of this dilemma dramatically undermines the suitability of Aim Accounts as the correct way to understand the truth-belief relationship.

In terms of those beliefs I was going to believe anyway, again appealing to a broken aiming mechanism doesn't appear to do Aim Accounts any favours.

As with the other *idées fixe* above, this type of *idée fixe* also follows other doxastic norms, putting Aim Accounts back on the horns of the same dilemma noted above.

Finally, the case of finding an argument both unanswerable and yet unbelievable probably has the best chance of being saved by an appeal to a broken aiming mechanism. Let's go back to Ribeiro's (2002) case to see why this defence may work here. Ribeiro, in the face of an argument for external world scepticism that he finds unanswerable, thinks his belief ought to change and ought to change, because, as far as he's concerned, his lingering belief is not based on epistemic reasons. This highlights that as far as Ribeiro is concerned, his belief fails other doxastic norms. As he argues, cogently, if I can find no epistemic reason for why my belief maintains, I must conclude that my belief is sustained for non-epistemic reasons (Ribeiro, 2002:471). In this instance, not only is Ribeiro unable to tell an epistemically rationalising story about his/her belief, he is also pretty convinced that there is no epistemically valid reason for his belief. The only thing that these *idées fixe* have going for them in terms of belief is that they are propositions that are taken as True in the sense that means a belief.

Given the veracity of this, then perhaps the way to understand beliefs that persist in the face of unanswerable arguments is that the mechanism that does the aiming for the belief is broken. This broken mechanism would explain why it's the case that we can

see that there's something odd going here, without our belief being moved in any way.

While this is a plausible explanation, I have some concerns. These concerns don't, in and of themselves, undermine the possibility that this explanation of the *idée fixe* is correct, but might suggest that we look for a better explanation of what's going on in these cases. (Something I try to do in Chapter 5.)

If a broken aiming mechanism is the reason that explains these *idées fixe*, then it is worth highlighting that I, for example, have a number of beliefs that are quite fundamental to my daily life, which according to this explanation must now be described as broken beliefs.

My beliefs in motion, in the external world, in free will and in the usefulness of induction are all beliefs I hold in the face of arguments for which I have no epistemically good response and must, therefore, be described as broken beliefs. While the centrality of these beliefs is of no significance, given the validity of an Aim Account, it does strike me as odd that so many of my fundamental beliefs would appear to have broken aiming mechanisms.

This is by no means a solid case against the broken-mechanism defence of Aim Accounts, but it's something that's definitely intrigued me. Surely, there should be a better explanation for why so many of my central beliefs cannot be shifted? In Chapter 5 I offer a different explanation of these beliefs.

In this section I considered two ways that Aim Accounts could deal with an *idée fixe*. The first was to suggest that other non-epistemic commitments could destabilise a belief's aim. I argued that while this might account for some *idées fixe* some of the time, the move was too implausible to account for all *idées fixe*. The second strategy was to consider whether *idées fixe* could be explained by a broken aiming mechanism. I argued that while this defence might explain why we find some arguments unanswerable and yet unbelievable, this move could not explain the other types of *idées fixe* I've examined.

7.2. The direct link between beliefs and truth

One of the requirements of an Aim Account is that the constitutive aim of belief does the work in terms of explaining the features of belief. Thus, in order to explain that beliefs regulate for truth, for example, we must be able to do so in virtue of the tension between a belief and the Truth. In the case of Aim Accounts that tension is cashed out in terms of beliefs aiming at the Truth.

In this section, I argue that by putting beliefs in this tension we are unable to explain the direct link between beliefs and truth. As I noted above, one of the features of beliefs is that there is a direct link between beliefs and what one takes as True. To believe that *p* is to immediately take it as True that *p*. In other words, what one takes to be True just is what we call a belief. It is this direct link between beliefs and truth that explains why it is the case that "[w]hen you change your mind about what is true, you have therein, *ipso facto*, changed your beliefs" (Hieronymi, 2006:53).

Now, if Aim Accounts are right, then what does the work is the gap between what we believe (truth) and the actual (Truth), with beliefs (what we think is True) *aiming* at the actual Truth. As Owens (2003) puts it, if beliefs aim at truth then beliefs are "necessarily adopted with the purpose of being true …" (Owens, 2003:285).[62]

If Owens' (2003) claim is right we form a belief not because it is True, but with the aim of it being True at *some later date*. And this is what I object to. What an Aim Account appears to commit us to is the following: in forming a belief that *p* I form a mental propositional attitude that *p* with the intention of *p* being True at some later state. I form a mental proposition, under this account that *aims* at being True.

But we don't believe something with the *aim* of it being True at some later date. We form a belief precisely because we already think that it *is* True. If Aim Accounts are correct, then there was a gap between what I believe and what I think might be True. Even if we remove talk of the believer here, then under an Aim Account a belief is

[62] I'm assuming that here Owens (2003) is using the term "purpose" as a synonym of "aim".

formed with the intention of being True in the future. This suggests that beliefs, which is what we take as being True are formed before the belief takes it as being True.

But this fails to account for the direct link between beliefs and truth. To believe that *p* is to take something as True (truth). But if Aim Accounts are correct, then beliefs are formed not because they feel they have "hit" the Truth, but because they want to or aim to in the future. In doing so, it also fails to account for what it is to be a belief. This direct link between beliefs and truth highlights what beliefs are. A failure to capture this feature, therefore, fails to capture what it's like to have a belief or be a belief.

Aim Accounts, therefore, fail for two reasons. The first is they cannot account for *idées fixe*. The second is that that they remove beliefs from truth such that the direct link between truth and belief is lost, which undermines Aim Accounts' ability to describe what it is to be a belief. I turn my attention now to other problems with Purpose Accounts of the truth-belief relationship.

8. Purpose Accounts and other problems

In this section I argue that much like Aim Accounts, Purpose Accounts put beliefs in the wrong tension with Truth, so that the whole teleological explanation falls apart.

I argued above that Aim Accounts suggest that beliefs are formed with the hope that at some point they will be True, which undermining what it is to have a belief. When we form a belief, we already think it's True. Thus, Aim Accounts in virtue of the explanatory tension create a gap between beliefs and truth, which subsequently undermines them as an account of the truth-belief relationship.

Similarly, I argue that Purpose Accounts put truth and belief into an untenable relationship with one another. The explanatory tension creates a situation, I argue, where Purpose Accounts must either postulate a gap between beliefs and truth, which means there is no direct link between belief and truth, or they must postulate that beliefs are infallible.

Let's go back to how Purpose Accounts understand the relationship between belief and truth. I argued above that in order for Purpose Accounts to explain how beliefs follow epistemic norms Purpose Accounts must have Truth, rather than truth, as the purpose of belief.

What does the explanatory work in terms of the features of belief in a Purpose Account is that beliefs have an inherent purpose, Truth. Thus, what explains the fact that beliefs respond to counter evidence, for example, taking only epistemic reasons into consideration, is that to do otherwise would be to fail the inherent purpose of belief.

I suggested that one way to understand this model would be to consider a trap that catches only unicorns. This trap would only spring once there was a unicorn in the trap. This trap would need to do certain things in virtue of this purpose. For example, it would need to ensure that it took only reasons to suggest a unicorn into account. If it took other reasons into account, it would catch things other than unicorns: a trap that failed to follow these norms could no longer be said to be a trap that catches unicorns, while a trap that catches a unicorn has successfully discharged its purpose.

Let's go back to belief. A successful belief is a belief that has fulfilled its purpose and caught the Truth, in much the same way that a successful unicorn trap is one that has caught a unicorn. It follows from this that a false belief, therefore, is an unsuccessful belief. These are beliefs that have failed to fulfil the purpose of belief and so are unsuccessful. But if beliefs can only be formed once they are True, then Purpose Accounts in their current formulation cannot account for error. Either they must suggest a gap between beliefs and truth in the way that I've suggested Aim Accounts are committed to, or they must claim that all beliefs are infallible.

To show that Purpose Accounts are on the horns of this new dilemma let's go back to the unicorn trap. Consider a trap that thinks there is a unicorn in it. The animal in the trap has fulfilled the criteria that the trap needs in order to differentiate unicorns from donkeys and is convinced, as a result of following the norms generated out of its purpose, that the animal is a unicorn. What do we expect to happen? It seems as though it must, in virtue of its purpose snap shut.

However, unbeknownst to the trap, what it's caught is a donkey. It may have been unsuccessful, but the trap has been sprung and so the purpose has been (unsuccessfully, this time) discharged. Given this, how do we make sense of the trap resetting itself when evidence comes to light that what it has is not a unicorn, but something else? The trap has sprung, so unless there's some other entity at play that's setting the traps, the trap would have no way of releasing what it takes to be a unicorn and no reason to think that what it has is not in fact a unicorn.

We have the same situation when we form a false belief. By our own lights the proposition in question fulfils all the correct criteria for a belief: we take it as True, we have what appear to be epistemic reasons for it, and so on. As a result of these reasons, a belief is formed. But, the belief is false.

How does this belief then regulate for truth unless there is something else that's setting the doxastic traps? I suggest there are two ways out of this problem; both of which put Purpose Accounts into trouble. The first is that the initial belief is a unicorn. Thus, the reason we have a snapped trap is because the belief is True. The second is that there must be some gap between beliefs and Truth such that when the trap snaps shut egregiously we're able to make sense of a belief then responding to counter evidence. I argue that the first puts us squarely on the horns of Shah's (2003) dilemma, while the second creates precisely the kind of gap I'm worried about in Aim Accounts.

If Shah (2003) used my unicorn trap example he would argue that if beliefs could catch donkeys, then beliefs admit to reasons other than epistemic reasons. Since we can only take epistemic reasons as reasons to believe our doxastic trap must accept only reasons that indicate unicorns. It follows, if we go this route, that if we form a belief at all then it can only be in virtue of the fact that it True. To mix analogies, if our belief has caught a unicorn, then it follows that all our beliefs must be cases where

our beliefs have caught unicorns. Following from this, all our beliefs must be True. Thus, our beliefs are infallible.[63]

This was, in part, Shah's (2003) problem with Purpose Accounts. If we have a purpose strong enough to account for transparency, we can't account for motivated beliefs. If my reasoning above is correct, it can't account for motivated beliefs or, in fact, for any kind of incorrect belief. To form any kind of belief it seems as though the inherent purpose of belief would need to be discharged. Once that purpose has been discharged then there can be no way for this to be re-discharged. The trap has been sprung and the belief has fulfilled its purpose. If we appeal to some other thing that resets the trap then we're no longer talking about the relationship between belief and Truth, but instead between belief, Truth and something else.

"But wait," the proponent of Purpose Accounts might say, "you're confusing truth with Truth. Obviously we don't have infallible beliefs. To have a belief is to take something as *being* True, but this doesn't mean that it *is* True. What Purpose Accounts claim is that the purpose of belief is the actual Truth. If we think about beliefs in terms of the trap you've mentioned, it's the gap between what we believe (donkeys) and the actual Truth (unicorns) that is the purpose of beliefs."

This brings us to our second option noting that there is a gap between what we believe and Truth. What happens if we adopt this gap? What explains a belief's regulation for truth under this model is the fact that there is an unfulfilled purpose that a belief is striving to discharge. In this picture, instead of having a candidate belief passing over the trap of belief, which springs shut with Truth and remains open with falsehood, we now have beliefs passing over a trap, which, to mix the metaphor slightly, tries to turn them into unicorns.

[63] We could of course make some exceptions that partially assist a defence of the Purpose Account. If the evidence that we have at our disposal is absolutely the best evidence we could possibly have at the stage that we're considering a proposition and nevertheless, unbeknownst to us, incorrect, then we could make sense of the trap springing egregiously and us forming a belief. In these cases we could make sense of doxastic error and still maintain that a Purpose Account with Truth as its purpose was correct. However, there will be cases when the evidence we have is poor or underdetermined as far the Truth-value of the proposition in question is concerned and yet we still form beliefs. As Owens (2002) notes, "people form beliefs on the basis of inconclusive evidence all the time …" (Owens, 2002:393). In these cases, when other evidence is readily available to us, it is hard to see how a Purpose Account with Truth as its purpose could account for us forming an incorrect belief.

But immediately we run into problems. The first problem is how do we make sense of the initial belief being formed? We can't appeal to the purpose of Truth. There must be some explanatory mechanism that accounts for the initial belief formation and then from there we could appeal to Truth doing its job. In addition, given what a belief is, the belief must both think that it has fulfilled the purpose of belief and simultaneously think that it hasn't fulfilled its purpose. This suggests that a belief isn't something we think is actually True, but something that, much like Aim Accounts, we form as part of the processes of getting to True cognitions. But if we accept this gap we fail to account for what it's like to be a belief. To believe something it to take some proposition as True it is to think one has a True cognition.[64]

[64] At this point one might want to suggest that beliefs have the purpose of knowledge. In this case we could, perhaps, make sense of forming a belief with the acknowledgement that though the belief takes the proposition in question as being True, it does not know it. Thus we form a belief and this belief has the purpose of knowledge, which makes sense of taking something as True and still being driven towards Truth in the form of knowledge. I mentioned in my introduction, that this way of cashing out the truth-belief relationship is something that I'm not going to address in this thesis as a result of time and space constraints. However I will briefly mention two objections here I have to this move. If beliefs have the purpose of knowledge, then we form a belief because we take it as True, but not as something we know. It seems to me that when I take something as True, I would also claim I know it. Should someone ask of a belief, "ah, but do you *know* it?" I suspect that if I thought about the proposition under scrutiny and claimed that I didn't in fact know it my belief would disappear as well – shifting to something like "I believe there's a good chance that *p* is True" rather than "*p* is True". Bernard Williams (1973) makes a similar point with regards to knowledge:

> [There] is a rather deep prejudice in philosophy, that knowledge must be at least as grand as belief, that what knowledge is, is belief plus quite a lot; in particular, belief together with truth and good reasons. This approach seems to me largely mistaken. It is encouraged by concentrating on a very particular situation which academic writings about knowledge are notably fond of, that which might be called the *examiner* situation: the situation in which I know that *p* is true, this other man has asserted that *p* is true and I ask the question whether this other man really knows it, or merely beliefs it. I am represented as checking on someone else's credentials for something about which I know already. That of course encourages the idea that knowledge is belief plus reasons and so forth … Consider the following situation. They are engaged in this illicit love affair; she comes running in, and says 'He knows'. Her lover says: 'You mean he believes it? Well we know it's true. Now has he got good reasons for what the thinks?' She says, 'I think he just picked it up from the gossip next door.' Would her lover then be right in saying 'In that case, he doesn't know'? This would be absurd … (Williams 1973, 146-147).

Williams' (1973) point seems similar to mine: that from a first-person, experiential perspective Truth and knowledge are indistinguishable. While knowledge accounts could appeal to the fact that knowledge just *is* more, I still struggle to see how Truth and knowledge differ. If beliefs aim at Truth (the objective Truth), what more is required for this to count as knowledge from the belief's perspective? My second objection is that as Edmund Gettier (1963) points out, even if we meet the conditions for knowledge we could get it wrong. So, if we assume that a belief feels that it has achieved the purpose of knowledge and still gets it wrong, then how do we account for this mental state's regulation for Truth/knowledge? Are we not now in the same situation of having a trap that needs to be reset? And, if so, in virtue of what is the trap being reset? Or do we simply claim that beliefs, while having the purpose of knowledge never actually achieve it? Shifting the purpose or target from Truth to knowledge I fear only postpones the challenges these types of accounts must face.

It also just seems odd to think that a trap that captures only True beliefs would take on board something that might be potentially True, or could be turned into something True, if it only captures True beliefs. Think of it this way. A trap that catches unicorns doesn't look at a horse and say "look, this might be a unicorn in the future, so I'm going to take it in as part of the process of catching unicorns and try to get this horse to become a unicorn at some point". A trap that catches unicorns, captures unicorns.

Thus, if Purpose Accounts want to get off the horns of Shah's (2003) dilemma they seemingly land themselves on another one: either they must accept that there is a gap between beliefs and truth, which is incommensurate with being a belief, or they must claim that all our beliefs are actually True. Since neither of these is a viable option, we have good reasons to abandon Purpose Accounts in favour of something else.

9. Conclusion

In this chapter I examined teleological accounts of the truth-belief relationship. I argued that there is a crucial distinction missing from the literature between Aim Accounts and Purpose Accounts. In the former beliefs aim at Truth, while in the latter beliefs have Truth as their aim. I argued that the failure to acknowledge this distinction has made the job of presenting a coherent teleological account, which is able to make sense of the features of belief, much more difficult.

Having made note of this distinction – as well as some of the pitfalls of failing to acknowledge this distinction – I sketched out what I think the two different teleological models might look like and examined how they would work.

I then examined two objections to teleological accounts in the literature: Owens' (2003) claim that beliefs cannot aim and a version of Shah's (2003) objection that teleological accounts either fail to account for transparency and transparency-like features of belief, or they fail to account for motivated beliefs. I argued that while Aim Accounts are deeply threatened by Owens' objection, they are not threatened by Shah's dilemma. Purpose Accounts, however, have the exact opposite worry. While

these accounts are untroubled by Owens' objection, Shah's (2003) teleologist dilemma (Shah, 2003:460-465) seem to completely undermine Purpose Accounts.

The fact that the two accounts admit to different threats reiterates the need to separate and spell out these two different accounts of the truth-belief relationship. Trying to force them together ultimately does a disservice to any kind of teleological account, as well as anyone who objects to these types of accounts.

Having laid out these two objections I then examined further problems with these accounts. I argued that in both cases the teleological model requires a belief to do too much. By trying to account not only for belief formation, but how beliefs then subsequently respond to counter evidence we create a gap between beliefs and truth that fundamentally undermines what it is to be a belief. Thus, while I think that teleological accounts are misunderstood, I still think that the teleological account is the wrong way of understanding the truth-belief relationship.

The failure of these accounts has led some to think that perhaps the best way to understand the truth-belief relationship is normatively. In the following chapter I examine Normative Accounts of the truth-belief relationship.

Chapter 4: The normative account

1. Introduction

In Chapter 2 I made clear an argument in the literature that claims that an adequate account of the truth-belief relationship must, in virtue of that relationship, be able to explain the features of belief.

These features, which I spelt out in Chapter 2, include both those that appear to indicate a strong link between belief and truth as well as those that suggest a much weaker link between belief and truth.

The features of belief that suggest a strong link with truth are: the direct link between belief and truth; the fact that beliefs regulate for truth; that the only reasons we can take as reasons for belief are those that we think are epistemic reasons; that we give, in all but the most unusual of circumstances, an epistemically rationalising story about our beliefs; that because of the direct link with truth we can't believe at will; and that there is a direct link between epistemic reasons and our beliefs.

On the other hand are those features of belief that suggest that our beliefs don't always track the truth in a way that the previous features of belief might indicate they do. Those features include: motivated beliefs; biased beliefs; false beliefs; beliefs based on inconclusive evidence and evidence-resistant beliefs (*idées fixe*).

I argued that a full account of the truth-belief relationship needs to be able to explain both kinds of feature if it is to count as an adequate cashing out of the truth-belief relationship.

In addition, in Chapter 2, I argued that there are several important distinctions missing from the literature; two of which that we need to make note of in this chapter. The first is a distinction between truth and Truth. The former is what we take as being True. The latter is what is actually True. This distinction acknowledges the potential gap between what we believe and what is actually the case. This distinction is important because the moment we talk about beliefs aiming at truth we need to ask

whether we're talking about beliefs aiming at what we take as True (truth), or at what is actually True. The second is a distinction between the relationship that a belief has with truth and the types of projects that an agent can engage in with regard to a belief or beliefs. I argued that projects like Pascal's Wager, for example, are projects that a believer can engage in. These projects, while they may highlight features of belief, tell us nothing about the truth-belief relationship in and of themselves.

Having laid out the features of belief and their significance with regards to an account of the truth-belief relationship, I then considered the suitability of teleological accounts as accounts of the truth-belief relationship in the last chapter. I expanded on the existing literature by noting that teleological accounts can and should be separated into two distinct theses: Aim Accounts and Purpose Accounts. David Owens (2003) in "Does Belief Have an Aim?" appears to attack the former notion, while David Velleman's (2000) "On the Aim of Belief" appears to defend the latter model. This missing distinction, I argued, has muddied both the ability to explicate a teleological model of the truth-belief relationship, as well as the criticisms of such models.

However, despite this distinction, I argued that neither of these accounts is able to make sense of the truth-belief relationship. Aim Accounts fail to make sense of both *idées fixe* as well as failing to account for the direct nature of the relationship between beliefs and truth.

Purpose Accounts, meanwhile, are trapped on the horns of two dilemmas. The first dilemma, by Nishi Shah (2003), points out that Purpose Accounts must either make the relationship between truth and belief so strong so as to fail to account for instances of wishful thinking, or so weak as to fail to account for the direct link between truth and belief (Shah, 2003:460-465). The second dilemma I argued that Purpose Accounts are trapped in is that they must either claim that beliefs can't regulate for truth, or, in order to account for this feature, must postulate precisely the type of gap that I feel undermines Aim Accounts.

I argued that irrespective of whether we talk about aims or purposes in both cases teleological accounts fail to capture crucial features of belief and so should be discarded as the right way to understand the truth-belief relationship. The failure of

teleological accounts, whether they are Aim Accounts or Purpose Accounts, means we need to examine other ways the truth-belief relationship can be cashed out.

In this chapter I examine normative understandings of the truth-belief relationship. This account sees the relationship between belief and truth as essentially normative such that the features of beliefs are explained by a normative link between beliefs and truth. Notable proponents of this understanding of the aim-at-truth metaphor include Shah (2003), Shah and Velleman (2005), and Ralph Wedgwood (2002).

It is important to note that discussions in the literature around Normative Accounts of the truth-belief relationship focus primarily on deliberation.[65] As I argued in Chapter 2 deliberation is properly understood as a project of the believer. Any account that focuses solely on deliberation, therefore, is an account only of deliberation and can not count as an account of the truth-belief relationship. Because of this, the Normative Account I present in Section 2 is a version of Shah and Velleman's (2005) argument.

My version of the Normative Account looks similar to an Aim Account. The concern, therefore, is that it falls prey to the same objections that I levelled at Aim Accounts in Chapter 3. However, in Section 3 I explain how Normative Accounts escape the criticisms levelled at Aim Accounts.

In Section 4 I examine three objections to this account put forward by Asbjørn Steglich-Petersen (2006). I argue that only his first objection is, strictly speaking, an objection to a Normative Account of the truth-belief relationship. The other two objections, as they stand, appear to be objections to normative understandings of deliberation rather than objections to a Normative Account of beliefs.

In Section 5 I examine a potential response to what I take to be Steglich-Petersen's (2006) only objection to a Normative Account; drawing on a clue from Allan Gibbard's (2005) work.

[65] Nishi Shah and David Velleman (2005), for example, argue that their normative explanation of the truth-belief relationship applies only to doxastic deliberation. They argue that a teleological explanation of the truth-belief relationship is perfectly suited to being able to explain other types of belief formation. In this chapter I explore a Normative Account of the truth-belief relationship, *simpliciter*.

In Section 6 I examine whether a Normative Account can explain the features of belief. My primary concerns are the ability of Normative Accounts to explain how beliefs are formed on inconclusive evidence and their ability to explain *idées fixe*. I ultimately conclude that a Normative Account should be discarded in favour of something else.

2. Transparency and the case for the Normative Account

In this section I lay out a version of Shah (2003) and Shah and Velleman's (2005) Normative Account. As mentioned above, it is a version, rather than a direct explication of their account in virtue of the fact that they argue that a Normative Account applies only to deliberation. However, as I've argued in Chapter 2 these accounts would then count as normative understandings of deliberation, rather than as Normative Accounts of the truth-belief relationship.

I begin by focussing on Shah and Velleman's (2005) account (which draws heavily on Shah's 2003 paper). I have chosen this account as it appears to be the first and most-cited instantiation of a fully-fledged normative understanding of something like the relationship between belief and truth. While Wedgwood (2002) thinks we need to understand the aim-at-truth metaphor normatively he presents no account of how this understanding of the metaphor explains the features of beliefs I discuss in Chapter 2. Therefore, this chapter deals closely with Shah's and Shah and Velleman's Normative Accounts.

As this account is developed out of what these authors identify as a problem for teleological accounts it is necessary to briefly retread some ground covered in Chapters 2 and 3.

Shah (2003) and Shah and Velleman (2005) argue that there is a feature of belief called transparency.[66] Having highlighted transparency, these authors argue that

[66] As I argued in Chapter 2 there are three ways transparency can be understood: as the direct link between beliefs and truth; as the claim that only epistemic reasons count as reasons for a belief; and/or as the claim that these epistemic reasons have a direct impact on our belief. I argued in Chapter 2 that all three of these understandings of transparency appear to be genuine features of belief, but that, however transparency is understood, this feature of belief must apply to all beliefs, rather than just

teleological accounts are trapped on the horns of a dilemma. Either teleological accounts, in virtue of postulating a teleological relationship between beliefs and truth, can't explain transparency, or they can't explain motivated beliefs.

As I argued in Chapter 3, Aim Accounts are seemingly immune to this dilemma. For Aim Accounts beliefs literally aim at Truth and it is possible, as Bernard Williams (2002) notes, for these beliefs to mistake some non-epistemic reason as an epistemic one; thereby preserving the transparency from the belief's or agent's perspective (Williams, 2002:83), but still providing an explanation for motivated beliefs. So while Aim Accounts fail for other reasons they are not caught on the horns of Shah's "Teleologist's Dilemma" (Shah, 2003:460).

The same response is not available to Purpose Accounts, however. By cashing out the truth-belief relationship as one in which belief has an inherent purpose, Purpose Accounts are forced into this dilemma. The relationship with truth must be in-built and so the relationship between belief and truth must tread a very thin line: weak enough to account for instances of wishful thinking, but strong enough to account for transparency and transparency-like features of belief. It's a line that turns out to be an invisible line.

So, while Aim Accounts can cite something like misguided target acquisition, no such option is available to Purpose Accounts. Shah's (2003) objection to teleological accounts, therefore, seemingly only applies to Purpose Accounts.

However, while Shah's (2003) Normative Account is born out of failures of Purpose Accounts, rather than a failure of Aim Accounts, given the other problems with both Aim and Purpose Accounts, Normative Accounts do offer us a potentially valuable alternative account of how to make sense of the truth-belief relationship.

Normative Accounts, Shah (2003) and Shah and Velleman (2005) argue, are able to explain transparency in virtue of the fact that beliefs have a standard of correctness

---

deliberative beliefs. Since the meaning of transparency is ambiguous, in what follows I'm using transparency to cover all three of these features of belief; differentiating between them where I feel it is necessary.

(SoC). This SoC is the observation that "a belief is correct if and only if the proposition believed is true" (Wedgwood, 2002:267).[67] It is this SoC that does the work. In order to follow this norm, Normative Accounts argue, we must follow the epistemic norms we've come across in the first two chapters. In other words, in order to ensure that a belief meets the SoC we would have to only take epistemic reasons into account when believing; our beliefs would have to regulate for truth; and so on.

The first claim that needs to be defended is that the SoC is a norm that could/must be followed. As it stands the SoC seems redundant. Given that a belief is "the attitude we have … whenever we … regard [something] … as true" (Schwitzgebel, 2010), then it must be the case that in order for a belief to be True, it must actually be True. However, far from being a trivial or redundant descriptive claim, so the above authors argue, the claim is in an important normative claim about beliefs. As Wedgwood (2002) explains:

> To say that a mental state is "correct" is to say that in having that mental state, one has got things "right"; one's mental state is "appropriate". To say that a mental state is "incorrect" is to say that in having that mental state, one has got things "wrong" or made a "mistake"; one's mental state is in a sense "defective". Clearly there is nothing wrong or defective about false propositions as such; what is defective is *believing* such false propositions. (Wedgwood, 2002:276)

Further underlining this as a normative claim, says Wedgwood (2002), is that it would be irrational to form a belief that went against the SoC (Wedgwood, 2002:268). Doing so would involve "having an incoherent set of mental states" (*ibid*.). We would know that to believe correctly we would need to meet the SoC, but nevertheless we would form a belief that goes against the SoC. The claim here is not that it is possible to

[67] Shah's (2003) formulation of the SoC is slightly different, "I argue that … the standard of correctness for belief … is expressed in the *prescription* to believe that *p* only if *p* is true …" (Shah, 2003:448), while Shah and Velleman's (2005) formulation matches Ralph Wedgwood's (2002) formulation: "We claim that conceiving of an attitude as a belief … entails conceiving of it as an acceptance that is regulated for truth, while also applying to it the standard of being correct if and only if it is true" (Shah &Velleman, 2005:497).

engage in this kind of project, but that even if it was, it would nevertheless be something we ought not to do.

As it stands, the claim is that the SoC for belief appears to pick out a normative claim about belief. The SoC provides a standard to which beliefs should adhere to, thereby allowing one to make normative judgements about those beliefs – both from a first-person and a third-person perspective. Beliefs that fail to meet the SoC are incorrect and ought not to be believed, whereas those that meet the SoC are correct and ought to be believed.

Now what happens in doxastic deliberation, argue Shah (2003) and Shah and Velleman (2005), is that the deliberative process is governed by the norm such that the feature of transparency is explained. In other words, because of the SoC, deliberative questions about whether to believe that *p* are transparent to questions about whether *p*.

Steglich-Petersen (2006) explains:

> Thus, the explanation goes, deliberation about whether to believe that *p* is transparent to deliberation about whether *p* is the case, because there is a *constitutive* norm for believing, saying that the correctness of believing *p* ought to be decided only by whether *p* is true, and one cannot decide to believe without deciding to care about this norm. In effect, one cannot decide whether to believe that *p* without deciding whether *p* is the case. To frame a deliberative question in terms of belief *just is* to accept the norm that beliefs are correct only if true; whatever one is motivated to do by acceptance of that norm is in virtue of one's understanding of the concept of belief. (Steglich-Petersen, 2006:506)[68]

[68] It is worth noting that not everyone takes the SoC for belief to articulate a *constitutive* norm of belief. Timothy Chan (2013) notes that while several people have felt that the SoC for belief is a constitutive norm of belief, several people find this claim questionable (Chan, 2013:3-4). I, along with Allan Gibbard (2005), think that the SoC for belief is entailed by, rather than constitutive of, the concept of belief. However, the reason I ignore this debate is as follows. I take it that the SoC articulates a necessary claim about belief such that, following Kit Fine's (1994:1994a) modal argument, we would see the SoC as a property of belief in all possible worlds. If we accept that it's a norm, as per Wedgwood's (2002) argument above, then is a necessary norm of belief. What would be at issue is

To follow the SoC would be to take only epistemic reasons as reasons to believe. If I were to take other reasons into account I could no longer be said to be trying to adhere to the SoC. Similarly, if I see that my belief is undermined by counter evidence I must, if I'm adhering to the norm of correctness, give up that belief and so on. As a result, transparency and transparency-like features of belief are explained in virtue of this norm of Truth.

A Normative Account, therefore, appears to do more than just explain why transparency is highlighted in cases of doxastic deliberation. In addition, this account seems to explain why beliefs would regulate for truth; the primacy and motivational force of epistemic reasons; the reason I can't give a pragmatic explanation for my belief (assuming I'm sincere and still hold the belief); and why we can't just choose to believe.

In addition, since beliefs in this instance attempt to adhere to the norm of correctness, rather than have Truth as their purpose (as is the case in a Purpose Account), we may be able to make sense of motivated beliefs. To believe that *p* is to think that *p* has met the norm of correctness. However, this doesn't mean the belief actually has met the SoC, only that it ought to.

> It is now clear how the normative account avoids the so-called teleologian's dilemma. The normative condition that beliefs are correct if and only if true, and the explanation of transparency in terms of motivation stemming from accepting this norm in applying the concept of belief to oneself, are compatible with weak *de facto* truth-regulation, in as much as a normative condition merely states how beliefs ought to be, not how they actually are. (Steglich-Petersen, 2006:506)

---

whether or not the SoC for belief would need to be a constitutive norm to do the regulatory work that a Normative Account relies on. What I'm not clear on is whether this regulatory work couldn't be done by a necessary, but merely entailed norm. Because of this I'm happy to accept that the norm could be efficacious. My argument is that Normative Accounts fail for other reasons. Should it transpire that the norm has to be constitutive in order to do the regulatory work in a Normative Account, this would then represent another challenge that proponents of Normative Accounts would need to address.

If you still find this as confusing as I do, this is how I make sense of it. There is a feature of belief, transparency. What is transparency? Asking whether to believe that *p* is transparent to asking whether *p*. What this means is that asking whether to believe that *p* is the same thing as asking whether *p* is True, or whether *p*.

What explains this transparency? According to Shah (2003) and Shah and Velleman (2005) teleological accounts (and they mean Purpose Accounts to my mind) cannot explain this feature of belief. Given that we're talking about purposes either the notion of truth must be weak enough to account for motivated beliefs – thereby ruling out the possibility of transparency – or it must be strong enough to account for transparency – thereby ruling out the possibility of motivated beliefs.

So what does account for transparency? According to these authors, what explains transparency is a norm that a belief is correct *iff* it is True. But how does this norm explain transparency? Well, if a belief is only correct when it is True, then the only way to try to get a True belief is to take into account evidence that speaks to the Truth of whatever claim we're deliberating about believing. Thus, in trying to determine whether to believe that *p*, one needs to determine whether *p* is True, or whether *p*. Thus, the deliberative question whether to believe that *p* is transparent to the question whether *p* is True, or simply, whether *p*. This is how the Normative Account explains transparency as I understand it.

Put like this, a Normative Account is able to make sense of the direct link between beliefs and truth – one of the ways we can understand transparency. One ought only to believe what is True or what one takes as True. In addition, Normative Accounts can also make sense of the other two ways we can understand transparency. If believing that *p* is governed by a norm to believe that *p iff p* is True, then believing that *p* requires one to take only Truth-directed reasons into account. If we take other reasons into account, then we fail the norm. Given this norm and given that I can only take epistemic reasons as reasons to believe, these reasons would have a direct impact on my beliefs – either causing us to form a belief in line with sufficient evidence or causing us to give up a belief in the face of sufficient evidence. Thus, no matter how we understand transparency a Normative Account seems well placed to make sense of it.

In addition to transparency, a Normative Account explains why beliefs regulate for truth: a failure to regulate in the face of sufficient evidence against the belief would mean that the belief is no longer following the norm. Thus, in order to ensure that it is in line with the norm, beliefs must regulate for truth.

So, one horn of the dilemma dealt with. How does this account deal with motivated beliefs? In order to explain transparency the Normative Account has to rule out the non-epistemic reasons that account for instances of wishful thinking. But, in saying that only epistemic reasons count as reasons in doxastic deliberation, the Normative Account looks like it's saying that doxastic deliberation is never wrong. And this can't be right. As Steglich-Petersen (2006) points out there are a number of different ways that we can get it wrong in cases of doxastic deliberation:

> What the deliberator *takes* to be relevant for the truth of the proposition under consideration might not be relevant at all, or not to the degree the deliberator takes it to be, and the deliberator might be swayed in evaluation of evidence without realizing this. It is only from the deliberator's perspective that the belief-formation is strongly regulated for truth. (Steglich-Petersen, 2006:511)

This, seemingly, places the Normative Account on the horns of its own dilemma: either the normativist must acknowledge that non-epistemic factors can influence belief-formation, which appears to undermine transparency, or he/she must deny that non-epistemic factors ever play a role in deliberation, which undermines that we could ever have a deliberative belief that is based on non-epistemic or irrelevant reasons.

The problem for the normativist is this: to accept the first horn of The Normativist Dilemma is to admit that non-epistemic factors can and do impact on belief-formation in deliberation. As a result, the normativist is able to account for instances of doxastic deliberation that have taken non-epistemic reasons as epistemic ones. However, in doing so the normativist is now seemingly trapped in precisely the dilemma levelled at the teleologist. If the Normative Account allows for non-epistemic reasons then how we make sense of transparency? The only other option is deny that non-epistemic

factors ever play a role in deliberation. However, as mentioned in Chapter 2 there is plenty of evidence to show that non-epistemic factors do play a role in all types of belief-formation.[69]

So how does a Normative Account succeed where a teleological one fails? The answer to this potential problem for the Normative Account lies in the distinction between Aim Accounts and Purpose Accounts I made in Chapter 3. I argued in Chapter 3 that Shah's (2003) dilemma is a critique of Purpose Accounts and fails as a critique of Aim Accounts. The Normativist Dilemma, therefore, is only a dilemma if Normative Accounts are like Purpose Accounts.

However, as laid out above, Normative Accounts, in virtue of intentionally following a norm, appear to be similar to Aim Accounts and are thus are able to escape The Normativist's Dilemma for the same reasons that Aim Accounts manage to escape Shah's (2003) dilemma.

As mentioned above, Aim Accounts can admit to non-epistemic influences in a way that Purpose Accounts cannot. In the case of Aim Accounts, the belief has simply missed its mark. The Normative Account, much like Aim Accounts, can appeal to the fact that in following the norm agents sometimes mistake a non-epistemic reason for an epistemic one. Transparency is saved, because from the agent's perspective the reasons are seen as Truth-directed reasons, while error is accounted for in virtue of the fact that non-epistemic reasons can creep into the deliberative process because they are mistakenly taken as epistemic reasons.

As Steglich-Petersen (2006) notes above, Normative Accounts "are compatible with weak *de facto* truth-regulation, in as much as a normative condition merely states how beliefs ought to be, not how they actually are" (Steglich-Petersen, 2006:506).

Thus, a Normative Account is able to explain several features of belief. It's able to explain the fact that the only reasons that count as reasons for belief are epistemic

[69] For examples of seeming deliberative beliefs that nevertheless appear to be motivated by non-epistemic factors see Kwame Appiah's (1990) example of Professor William Shockley and Erik Nisbet and Kelly Garrett's (2010) research on Americans who believe that US President Barack Obama is not an American citizen despite being exposed to objective evidence to the contrary.

reasons; it's able to account for the fact that our beliefs would regulate for truth; it's able to account for the fact that epistemic reasons would speak directly to the beliefs those reasons impacted; and it would also be able to account for motivated beliefs. Further, my suggestion is that this account be understood, not as an account of deliberation, but rather as a fully fledged account of the truth-belief relationship.

I argued in Chapter 2 that if transparency is to count as a feature of any kind of belief, then it must be a feature of all beliefs. Otherwise it's not a feature of belief, but rather a feature of the project the agent is engaged in.

Similarly, we must be careful to distinguish between an account of deliberation and an account of the truth-belief relationship. I argued in Chapter 2 that there is a distinction that needs to drawn between the project of a believer and the inherent relationship that belief has with truth. I argued that while doxastic projects such as deliberation or Pascal's Wager can highlight features of belief, we shouldn't confuse the norms generated out of a project of the believer's with the ability of an account of the truth-belief relationship to explain why beliefs follow epistemic norms.

As Ward Jones (2002) points out, if believing is "a goal-oriented process … it is not … so in virtue of desires or intentions that the believer has" (Jones, 2002:222). Thus, if believing is to be understood teleologically, then even if a believer isn't interested in aiming at the Truth (as appears to be the case in Pascal's Wager) this belief would, in virtue of the relationship beliefs have with truth, aim at Truth.[70] The belief would aim at Truth irrespective of the project of the believer because of how the relationship between beliefs and truth is constituted. If beliefs aim at Truth then they do so independently of whether or not an agent aims at Truth. Similarly, if the relationship between beliefs and truth is a norm-following one it is so irrespective of whether or not the agent wants to follow that norm. What this means is that in order for a Normative Account to count as an explanation of the truth-belief relationship it must apply to all beliefs, rather than just deliberative beliefs. Should it apply only to

[70] Pascal's Wager is the argument that we should try to believe in God given the prudential benefits of such a belief. When I claim that Pascal's Wager doesn't aim at the Truth, the claim I'm making is not that a belief is God is false, but rather than in Pascal's Wager the aim is to believe that God exists irrespective of whether or not it might be True. This means that if one is engaged in Pascal's Wager we're not aimed at Truth, but rather at a specific belief.

deliberative beliefs, then it is an account of how an agent should engage in the project of deliberation rather than an account of the truth-belief relationship.

Happily for Normative Accounts, I can't see, given what's been said, why this should be problematic. Given that the feature that Shah (2003) and Shah and Velleman (2005) are trying to explain is a feature of all beliefs this suggests that the Normative Account is an account of the truth-belief relationship rather than an account of deliberation. Furthermore, what stands in favour of this as an account of the truth-belief relationship is that this account explains more than just transparency: it also makes sense of why beliefs regulate for truth as well as explains motivated believing.

In addition, suggesting that beliefs are governed by a constitutive norm seems no stranger than thinking that beliefs are governed by a constitutive aim (such that beliefs aim at Truth in Aim Accounts), or that beliefs have Truth as their aim (Purpose Accounts).

However, despite the appeal of a Normative Account given its ability to explain those features of belief that Purpose Accounts, in particular, fail to make sense of, there are several objections to this account. I start by dismissing one potential objection, before examining three objections to this account by Steglich-Petersen (2006).

3. Escaping the problems of an Aim Account

I've argued above that Normative Accounts are able to escape Shah's (2003) dilemma in much the same way that Aim Accounts do. Given this solution, surely Normative Accounts are susceptible to the same problems that face Aim Accounts?

I argued in Chapter 3 that Aim Accounts fail because they place belief in the wrong relationship to truth. Is the Normative Account not prone to precisely the same objection? The short answer is no.

The longer answer goes like this. Aim Accounts fail because they are forced to claim that beliefs are formed before what the agent takes as Truth is established. But, as I noted in Chapter 2, there is a direct link between beliefs and truth. The claim that

beliefs aim at Truth suggest that a belief is formed and then aims at the Truth. However, as I argued, this undermines the claim that to believe that *p* is to already take *p* as being True. Since this is the very definition of belief, Aim Accounts fail by placing belief in the wrong relationship with truth.

Under the Normative Account, however, rather than the belief doing the work of aiming at Truth, it is the norms that guide the belief, rather than the belief itself. We could say that the norms aim at Truth, rather than the belief in and of itself aiming at Truth. And the beliefs are formed when the norm has been adhered to; or thought to have been adhered to. That said, unlike Purpose Accounts, which as I argued in the previous chapter might suggest that we can only form a belief when it is in fact True, Normative Accounts make sense of error by acknowledging that beliefs might sometimes think they have met the SoC for belief; when in fact they haven't.

What this does is rather neatly avoid the pitfalls of Aim Accounts and simultaneously avoid the problems inherent in Purpose Accounts. Further, while Shah (2003) and Shah and Velleman (2005) think that a Normative Account applies only to doxastic deliberation, as I argued in Chapter 2, if transparency is a feature of belief at all, then it is a feature of all beliefs. Thus, if transparency is a feature of all beliefs and a Normative Account is able to make sense of it and motivated beliefs, and avoids the problems of an Aim Account, then it seems well placed to explain all the features of belief. The Normative Account, therefore, seems like a good account of the truth-belief relationship.

However, this is by no means the only objection to the Normative Account. In the following section I examine three objections put forward by Steglich-Petersen (2006).

4. Steglich-Petersen and the Normative Account

Steglich-Petersen (2006) argues that several key features of belief are not adequately explained by a norm-following account.

I argue that of the three concerns Steglich-Petersen (2006) puts forward, only the first is an objection to the Normative Account. The other two are objections to a normative

understanding of deliberation. In what follows, I examine each of Steglich-Petersen's (2006) objections as objections to a Normative Account, exploring each for its suitability as an objection the Normative Account of the truth-belief relationship. I argue that only one of these objections applies to the Normative Account.

Steglich-Petersen's (2006) first objection is that if we cash out the link between belief and truth normatively the epistemic norms associated with belief could not motivate us to follow them in the way that we see them do (Steglich-Petersen, 2006:506-507). I argued in Chapters 2 and 3 that what makes these doxastic norms interesting is that our beliefs follow them often unconsciously, sometimes in the face of our desires for the world to be otherwise. Any account of the truth-belief relationship, therefore, has to be able to explain not only transparency, but the motivational force of the epistemic norms on our beliefs. If a Normative Account cannot explain how it is that these norms motivate us to act, then a Normative Account should not be considered a good theory of the truth-belief relationship.

> While it is plausible to suppose that there are certain concepts that one cannot apply without accepting certain evaluations, it is not plausible that such evaluations can necessarily motivate us to act in accordance with them. (Steglich-Petersen, 2006:506)

So, while we can make sense of something having a norm that applies to it, there seems to be nothing, conceptually, stopping us from disregarding the norm in question. But doxastic norms aren't things that we can choose to follow or not to follow; they are things that we follow seemingly automatically. That our beliefs regulate for truth, for example, seems outside of our voluntary control. As I've argued in Chapter 2, when you see that *p* is correct, no matter how much you'd like *p* not to be correct, you find that you can't help but believe that *p*. But given that other norms are norms that we can decide to adhere to or disregard, how do we explain this regulation for truth in virtue of the norm of truth?

Steglich-Petersen (2006) compares the SoC for belief with promising. He states that it seems reasonable to assume that it is a conceptual truth about promises that if you promise to do something then you ought to do it. As he puts it, "[t]his plausibly means

that one only counts as promising to Φ if one acknowledges that one ought subsequently to Φ" (Steglich-Petersen, 2006:506).

However, as Steglich-Petersen (2006) points out, in promising, while you can be motivated by the norm, there is an intermediary step between making a promise and deciding whether to fulfil the conceptual norm. We can decide that we don't want to fulfil the norm of keeping the promise. In this case the norm needn't motivate us. However, with belief, in order to account both for transparency and the subsequent regulation of beliefs for truth the norm *must* motivate us.

Transparency only works if the norm tells us we can only take epistemic reasons as reasons for belief *and* if we're compelled to act in accordance with those reasons. While transparency explains why we should only take a certain type of reason into consideration if we are to be considered to be engaged in believing, it doesn't explain that we do so automatically. This is similar to the norm generated out of promising. I can't actually be said to be promising if I know in advance that I'm not going to follow through with that promise. Thus, promising does explain why I should follow the norm. However, in promising, I can decide not to follow through with this norm. So why then can we not decide to disregard norms generated out of the SoC for belief? Surely, if believing is governed by a norm in much the same way that promising is, then I can choose to not only disregard the SoC for belief but in addition the concomitant norms that come along with it. But we can't decide to act against these norms.

In fact one of the key challenges of a truth-belief relationship is being able to explain why we can't disregard these epistemic norms. Thus, if Steglich-Petersen (2006) is correct, then a Normative Account cannot account for the motivational efficacy of the epistemic norms and should, therefore, be abandoned. As he puts it, given what we know about norms, "[it] is thus doubtful whether a consideration which necessitates motivation should be considered a normative consideration at all" (Steglich-Petersen, 2006:507).

This is a serious objection to the Normative Account. If, in virtue of being a norm, we are necessarily able to follow or disregard the SoC, then while a Normative Account

can explain why one ought to only take epistemic reasons as reasons for belief it can't explain why we in fact do so. The SoC for belief highlights that a correct belief must in fact be a True belief in order to be considered correct. Thus, if we're trying to believe correctly then there are certain epistemic norms we ought to follow. One of those is that only epistemic reasons for *p* ought to count as reasons to believe that *p*. So far, so good, the SoC for belief does explain why this epistemic norm ought to be followed. However, when we examine norm-following behaviour we see that it involves an intermediary step such that while we appreciate that there is a norm that we ought to follow we can always decide not to follow the norm. If this is right then in deciding whether to believe that *p*, while I acknowledge that I ought to take only epistemic reasons as reasons for the belief I could choose not to.

The Normative Account, therefore, *can* in fact explain why we can only take epistemic reasons for *p* as reasons to believe that *p*. However, this is cold comfort. What it can't do is explain why, as a matter of course, these epistemic reasons have the motivational force they do.

Steglich-Petersen's (2006) second objection focuses on why the deliberating agent tries to believe correctly (Steglich-Petersen, 2006:507-509). The way Shah and Velleman (2005) have cashed out the Normative Account, Steglich-Petersen argues, suggests that the reason an agent tries to believe in line with the SoC is not as a result of being interested in Truth, but instead is a result of the agent's desire to adhere to the norm.

The objection here works like this. If Shah and Velleman (2005) are right, then in deliberation we're not interested in Truth, but instead are interested in being right, or being correct. This seems weird. If we look at the promise example again it would be like saying that in fulfilling a promise we don't do it because we made a promise to something and so in fulfilling the promise we're doing the right thing, but rather we do it to make sure we fulfil the norm.

If this seems like too much of the same thing, consider different motivations for not killing someone. The first is because killing someone is against the law. The second is that killing someone would fill you with moral horror. Now, if people didn't kill each

other solely and always for the former reason we would feel like something had gone "wrong". In other words, this seems like the "wrong" motivational reason for not killing people. Similarly, it seems as though we don't try to believe correctly in order to fulfil the norm, but instead we try to believe correctly because the Truth of the proposition in question is important to us.

> When I consider whether or not to believe that the train leaves at 5 p.m., I am interested in arriving at the true answer because it matters to me whether or not it is true: I might otherwise miss the train and end up waiting for the next. My interest is *not* adherence to a norm, unless perhaps doing what is required by the norm results in something I am interested in, namely, the truth. But that would make the truth-norm instrumental, which is ruled out by Shah and Velleman; and in any case it could not explain the immediacy of transparency as they describe it. (Steglich-Petersen, 2006:508).

James and Stuart Rachels (2010) make a similar point in their discussion on virtue ethics. In critiquing Deontology they provide the following example:

> [Imagine you] are in the hospital recovering from a long illness. You are bored and restless, and so you are delighted when Smith comes to visit. You have a good time talking to him; his visit is just what you needed. After a while, you tell Smith how much you appreciate his coming – he really is a good friend to take the trouble to come see you. But, Smith says, he is merely doing is duty. At first you think he is only being modest, but the more you talk, the clear it becomes that he speaking the literal truth. He is not visiting you because he wants to or because he likes, but only because he thinks he should "do the right thing." He feels it is his duty to visit you ... (Rachels & Rachels, 2010:168).[71]

As Rachels and Rachels (2010) note, irrespective of whether it is done out of virtue or duty the outcome is the same, Smith has visited. The problem, they suggest, is that

[71] Rachels and Rachels (2010) note that this example is originally Michael Stocker's, though this formulation is their own.

Smith's motivation for the visit, duty, seems like the wrong reason altogether; it seems to miss the point of being moral.[72]

Similarly, in terms of believing, we arrive at a belief whether we follow the SoC for belief or out of an interest in finding out the Truth. However, in the case of the former it seems to completely miss what it is to deliberate. Or even what it is to believe.

Doxastic deliberation presupposes an agent's interest in the Truth of the proposition deliberated upon, and this is something that cannot be cashed out solely in terms of the norm suggested. If this is correct, then it seems that positing believing or deliberating as having the goal of truth coheres better with our experience of either than a norm-following description does.

As it stands, however, this objection would apply only to deliberation. Steglich-Petersen's (2006) contention is that deliberation, *qua* deliberation, is better explained by postulating an aim than it is by a constitutive norm. However, while this may be a valid claim about deliberation, this is not an objection to a Normative Account of the truth-belief relationship. It may well be, for example, that norm-following explains the internal mechanisms of how a belief is formed, but that having an aim of Truth explains the project of the believer in the case of deliberation. Similarly, in Pascal's Wager there is an aim to reach a certain belief. In this case explaining the believer's project in terms of aim-following behaviour makes sense, but this doesn't tell us anything about what the relationship is between *belief* and truth.

However, while I think that Steglich-Petersen's (2006) objection is primarily against deliberation rather than a Normative Account of the truth-belief relationship there is something a little odd about beliefs being formed in virtue of being interested in fulfilling the SoC rather than interested in the truth of the proposition in question.

There is an example of Barbara Winters (1979) where an agent "is truthfully informed that either a gentleman or a hungry tiger waits in the next room and that the outcome has been decided by the concealed flip of a fair coin" (Winters, 1979:250). It seems

[72] The aim here is not to defend theories of virtue over deontological theories. I merely find this a useful way of highlighting the weirdness that Steglich-Petersen (2006) is alluding to.

that should the agent incorrectly come to believe that there is a gentleman behind the door she enters, her belief will quickly regulate for truth without the need for deliberation.

If a Normative Account is correct, then the reason her belief regulates is not because there is any interest in truth, but because the belief is trying to discharge the norm; and this seems odd. It seems that in this case the interest would be in the truth of the matter, rather than getting it right for the sake of discharging the norm. However, this is by no means a knockdown objection to the Normative Account. It may simply be a quirk of our internal cognitive make-up that beliefs, *qua* beliefs, are regulated by a norm which they seek to discharge no matter how odd that might seem given our lived experience of beliefs.

Steglich-Petersen's (2006) third objection is that a constitutive norm cannot account for the fact that when we form deliberative beliefs we do so on different amounts of evidence (Steglich-Petersen, 2006:509). This objection is a corollary to the second objection. By cashing out truth-seeking behaviour in terms of an adherence to a norm rather than an interest in truth itself, what explains why it is that we sometimes require less evidence to form a deliberative belief? Steglich-Petersen (2006) notes that if we cash deliberation out in terms of a goal of Truth, then we can understand why sometimes the Truth is more important to us than at other times. However, if believing is simply about discharging the norm of correctness, then how do we explain the fact that we require different amounts of evidence in different cases?

> … what explains why we are willing to form a belief on less evidence on some occasions than on others? On the normative account, our interest in explicit belief-formation is in complying with a norm for belief, i.e., ending up with a belief which is correct according to the norm governing belief. But it cannot explain why we should be more interested in making sure that we comply with the norm on some occasions than on others: a true belief is correct no matter what its importance is, and it does not become more correct by being more important. (Steglich-Petersen, 2006:509)

The concern here is that it seems that if I am forming a deliberative belief about when an important job interview is, I'm going to put in a quite a lot of deliberative work in to ensure that I have the truth of it. Check, check and double check as my great-uncle used to say. However, if I'm trying to work out when the reality show *Survivor* re-runs I may simply check the TV guide and have done. If we understand belief formation as being inherently interested in Truth, rather than norm following, the answer is simple. Some answers/truths are more important to an agent than others. My interest in the time of the job interview, supersedes the rather lesser importance of what time the re-run of *Survivor* is on this weekend.

Explaining the differing levels of interest in truth helps us understand why, with regards to some propositions, a greater degree of evidence or certainty is required before the proposition is believed. In some cases the Truth is very important to us, while in others not so. The greater the interest in the Truth the more likely it is that we will try to establish the Truth of that proposition. Trying to cash this out in terms of following a norm, however, doesn't work quite as neatly. What explains the desire to follow the norm more in one instance than in another?

Having differing levels of interest in the truth of a proposition make sense if we treat doxastic deliberation as one goal among many other goals. Given the importance of the aim of Truth in some instances versus others, this explains why we might need more evidence in some cases than in others. But the normative constraint is a prescriptive requirement to believe that *p* only if *p* is True. This suggests that in each instance there ought to be the same level of evidence, or evidence gathering, required in each instance of deliberation to satisfy the norm.

However, this is clearly not the case, which challenges following a norm as the right way of understanding deliberation. However, while I think that Steglich-Petersen (2006) correctly highlights this as a problem for deliberation, it seems to me that differing interest in the Truth of propositions isn't, as it stands, a problem of the truth-belief relationship. While the SoC as it applies to deliberation may not be able to

explain why in some cases of deliberation I require less evidence than in others, this is an objection to how the norm operates in deliberation.[73]

It seems then that in terms of Steglich-Petersen's (2006) objections only the first one, as it stands, directly applies to a Normative Account. For if the relationship between beliefs and truth is understood normatively, how do we make sense of the fact that in all other norm-following behaviour we have an option to disregard the norm as well as the concomitant norms generated by that norm? In the following section I examine how a Normative Account could respond to this challenge.

5. A case for a different kind of norm

In this section I draw on a clue in Gibbard's (2005) work to try to defend the Normative Account from Steglich-Petersen's (2006) objection above. The claim I try to defend here is that there is a fundamental difference between the type of norm expressed by the SoC and other pragmatic and moral norms. I explore why we might think there is a case for a distinction between epistemic norms and "other" norms.

In "Truth And Correct Belief" Gibbard (2005) is trying to work out how we get the normative SoC from a factual claim about belief. The concern for Gibbard is, following David Hume (in Gibbard 2005:338), that we seem to be getting an "ought" from an "is". Just because a belief is true, doesn't mean it ought to be true. Similarly, just because a belief is false can't justify that it ought not to be this way (*ibid*.).

While getting the normative claim from the factual claim is an interesting question I'm not going to deal with here. I'm happy to accept for the purposes of this chapter that the SoC for belief is a normative claim about beliefs. What is interesting, for the purposes of this chapter, is that Gibbard (2005) points out that we need to be careful about comparing epistemic norms with moral or pragmatic norms.

[73] While this objection of Steglich-Petersen (2006) doesn't threaten a Normative Account as it stands, it does highlight that in instances of believing there are times when our beliefs appear to be satisfied by evidence that surely would not discharge the SoC. This would be an objection to a Normative Account. I consider this objection in Section 6 below.

Gibbard (2005) notes that we don't get moral or pragmatic notions of correctness from determining if a belief is true, all we can get is epistemic correctness (Gibbard, 2005:340).[74] This highlights that there appear to be two different types of norm at play here; what Gibbard (2005) calls "an important contrast in kinds of ought" (Gibbard, 2005:340). Now perhaps these different norms have different rules. If there is a valid distinction between "other" norms and epistemic norms, perhaps it is the case that epistemic norms simply don't allow the agent to engage with them in the same way that we engage with "other" norms. The argument against Steglich-Petersen's (2006) objection, therefore, is that there is something unique about epistemic norms such that unlike other norms we can't help but follow them.

But why should we think that there are two different types of norm here? One possible argument in favour of making this move is that given that epistemic reasons appear to be a different type of reason, it follows that epistemic norms are a different type of norm.

5.1. Epistemic reasons and a different type of norm

As we've already seen in Chapter 2 there is something odd about epistemic reasoning. While practical reasoning is able to respond to a variety of different types of reasons – self-interest, moral, economic, epistemic, etc. – for performing an action, doxastic reasoning responds solely to (what an agent takes as) truth-bearing reasons.

For example, as an instance of practical reasoning, I can decide to plagiarise my thesis despite there being overwhelmingly good reasons not to. I can't, however, believe, *ceteris paribus*, that my partner is faithful in the face of sufficient evidence to the contrary, no matter how much I want to.

What this highlights is that in practical reasoning I can bring a number of different types of reasons to bear on my decision: epistemic reasons, prudential reasons,

[74] One could argue that someone who has correctly reasoned to the conclusion that racism is wrong has reached a conclusion that is morally correct in virtue of his/her epistemic correctness. I would argue that in this case the agent has satisfied two norms: one moral, one epistemic. These two norms are fundamentally different; though, as in this case, they intersect with one another. Epistemic correctness pertains only to its truth condition, whereas moral correctness would pertain to whether or not the belief was morally acceptable. In this case, the belief is both epistemically correct and morally correct.

hedonistic reasons, moral reasons, and so on, some of which are Truth-directed in nature, some of which are patently not.

In addition, in practical reasoning I can at any point decide to ignore every good reason (however one makes sense of good) for *X* and decide to do *Y*; where the reason might be the infantile "just for the hell of it". However, epistemic reasoning doesn't work like this.

As Thomas Kelly (2003) points out, if epistemic reasons were like practical reasons we could choose whether to act on them or not. However, no such choice is available to us with regards to epistemic reasons. No matter how undesirable it is to believe that my partner is cheating on me I nevertheless find myself believing that conclusion when presented with sufficient evidence for that claim. In practical reasoning I can act "just for the hell of it". In epistemic reasoning my hand is forced by the evidence.

Now, Steglich-Petersen's (2006) first objection rests on the claim that in pragmatic reasoning there is an intermediary step in which one decides whether to act on those norms or not. Transparency can't allow for this. Transparency is a feature of belief. Therefore, beliefs are not controlled by following a norm. But given the unique nature of epistemic reasons might it not be the case that epistemic norms don't allow one to have an ability to decide whether to follow them or not?

The argument against Steglich-Petersen (2006) would work like this.

1. Epistemic reasons are not a type of practical reason and should be considered as a unique type of reason.
2. Given the uniqueness of epistemic reasons it seems reasonable to think that epistemic norms are different to "other" norms. In fact, when we talk about the motivational force of doxastic norms we note that these aren't things that we can choose to follow or not, we just follow them.
3. There seems to be nothing, conceptually, that rules out norms that can't be ignored.
4. If this is the case, then in much the same way that we can't help but follow epistemic reasons, perhaps we can't help but follow epistemic norms.

5. Therefore, instances of epistemic normativity needn't be similar to instances of pragmatic normativity. As a result, a Normative Account can respond to Steglich-Petersen's (2006) first objection by denying that the type of intermediary step involved in moral or pragmatic normativity is present in epistemic normativity.

However, the problem here is in the first premise. Steglich-Petersen (2006) might object that what makes epistemic reasons unique is precisely what's at issue. Why is it the case that we can't choose between epistemic reasons and other types of reasons? Teleological accounts can talk about the constitutive aim or purpose of belief as the thing that makes epistemic reasons in some sense unavoidable, while Normative Accounts, in virtue of appealing to following a norm, are seemingly unable to do this. Teleological accounts, therefore, can explain the difference between the types of norms in virtue of a constitutive aim or purpose, while Normative Accounts can't.

Thus, to appeal to epistemic reasons as a way of defending the claim that epistemic norms are just different, is question begging. In virtue of what are these norms different? Teleological accounts are able to provide an explanation for this difference that Normative Accounts cannot.

5.2. Epistemic norms *just are* different

The other option for Normative Accounts is to claim that while they can't say why epistemic norms are different, we notice as a matter of course that they are different. As Gibbard (2005) has pointed out in determining that a belief is true all we can get is epistemic correctness (Gibbard, 2005:340), and maybe this difference has something to do with the norm, rather than the way the truth-belief relationship governs epistemic norms. Let's see if the claim that epistemic norms just are different such that they intrinsically motivate us is defensible.

The claim being defended here is: perhaps it is possible to have a norm that we can't help but follow. The question is, is this consistent, or allowed by the concept of a norm? If not, then this move to defend Normative Accounts from the objection above fails, and a Normative Account should be abandoned.

The obvious objection to the claim that we can have intrinsically motivating norms is that traditionally normativity has operated under the slogan of "ought implies can".[75] That in order for something to count as a norm it must, at least in theory, be possible that the agent could either follow the norm or not follow the norm.

> The judgment that you shouldn't have done X implies that you should have done something else instead; that you should have done something else instead implies that there was something else for you to do; that there was something else for you to do implies that you were *able* to do something else … (Van Inwagen, 2009:254)

If a norm requires one to have been able to do otherwise, and doxastic norms are norms that we can't help but follow, then they are not norms at all. In much the same way we would be reluctant to call the law of gravity a norm, we could not call an intrinsically motivating doxastic norm a norm at all. In this case gravity isn't a normative constraint so much as a feature of the world we live in.

If this is right, then there is a conceptual constraint against there being norms that we can't help but follow.[76] What this suggests is that if epistemic norms are to be considered norms at all they must admit to the inferential step Steglich-Petersen (2006) is worried about. If this is right the Normative Account is damned either way it goes: either we are not talking about norms at all, in which case the account is clearly not a normative one, or the norms must involve the inferential step that rules out the very feature a Normative Account is meant to explain.

---

[75]Not everyone thinks that the ability to do otherwise is a necessary feature of normativity. Robert Adams (1985) argues that we can still be held accountable for certain things irrespective of whether or not we have or had any control over them, while Harry Frankfurt (1969) argues that we can make sense of moral responsibility in a determinist universe where one could not do otherwise. What I want to do in this section is determine whether this objection would threaten a Normative Account, even *if* the slogan "ought implies can" is correct.

[76] There may be some concern that if this is correct then teleological accounts are also in trouble. As we've seen teleological accounts do postulate epistemic norms that our beliefs are constrained to follow. If this is the case then if it is conceptually impossible to have an intrinsically motivating norm, then surely all accounts of the truth-belief relationship that talk about the constraint of epistemic norms are in trouble. The reason this is not a problem for teleological accounts is that the motivation of the norms in teleological accounts is not in virtue of the norm, but in virtue of the aim. If we follow the aim intrinsically, we can't help but follow the constitutive norms. The claim Normative Accounts seemingly have to rely on is that some norms in virtue of being a norm could intrinsically motivate us, and this is what is at issue here.

Now we need to be a little careful here. Even if we accept the slogan that "ought implies can" as a legitimate and necessary requirement for moral responsibility, there is a difference between cases described by this slogan and the doxastic cases we're examining.

Peter Van Inwagen (2009) argues that in holding an agent morally accountable for doing *X* he/she must have been able to avoid doing *X* (Van Inwagen, 2009:254). Thus, if some cruel, unbeknownst-to-you twist of metaphysical fate means that your inability to travel faster than the speed of light means the death of your dog (Fido) it seems inappropriate to hold you morally responsible for Fido's death – although we can, of course, hold you causally responsible. In this example, you can't do anything different. No matter what you try to do it's a physical impossibility to achieve the state required to avoid the moral wrong. It would seem wrong, therefore, to hold you morally responsible for the death of Fido.

In the case of the doxastic norm, however, the agent can do something different. He/she can believe something that the evidence objectively shows is false. What he/she can't do is *believe* that he/she has believed falsely. Thus, there is a distinction between a norm that you can't help but follow and a norm that you, *ceteris paribus*, can't help but *believe* you follow.

It's not that the norms governing belief *can't* be broken – we contravene these norms with frightening regularity. Instead we are constrained from believing that one of our present beliefs is both a belief and simultaneously one that fails to adhere to the relevant doxastic norm.

As we saw in Chapter 2 this type of constraint on what we can believe about our beliefs is hardly a novel suggestion. Winters (1979) and Jones (2002), for example, have argued that one cannot believe that *p* and believe that one's belief that *p* is sustained or explained non-epistemically, while Jonathan Adler (2002) and Michael Hicks and Adler (2013) have argued that it is impossible to believe that our belief that *p* is unsupported by adequate reasons for the truth of *p*.

While these constraints are not cashed out in terms of norm-following, the notion here is similar. Namely that the concept of belief imposes constraints on what we can believe about them. Thus, it seems within the realm of possibility that we may not be able to believe we contravene the SoC for belief; even though we can and do fail to meet it every time we form a false belief.

While this is undeniably odd, there doesn't seem to be any conceptual constraint that denies that we need to able to *believe* that we could have denied or followed the norm. The maxim, if this is the case, would be "ought implies that one believes that one can". So, in the case of the lethargic, dog-killer above, if you believe you could travel faster than the speed of light, then we would hold you responsible for Fido's death. But if you didn't we wouldn't. But that's not right. We don't find you innocent of murder under this slogan because you don't believe you can't travel at superluminal speeds, we find you innocent under this slogan because you *can't* travel faster than light. But if you could travel faster than light and didn't, then we can hold you responsible. So, while one might be constrained to think that one always acts in accordance with doxastic norms, this doesn't mean that we can't break them. We can, and do.

If this is correct, then it suggests that the inferential step that Steglich-Petersen (2006) correctly notes as being prevalent in most cases of norm-following might not be there in cases of epistemic-norm following. And the reason for this is that the epistemic norm (the standard of correctness for belief) simply doesn't allow one to think that one could not follow it and still think of oneself as believing.

But this doesn't mean that an agent can't break the norm. He/she can contravene the norm, so long as he/she believes that *p* and believes that the reasons he/she has for believing *p* are reasons that answer the question whether *p*.

If this is right, then Steglich-Petersen's (2006) first objection is potentially undermined. Steglich-Petersen (2006) was worried that in other norm-following cases there is an inferential step between accepting the norm and then deciding whether to follow that norm. However, I've argued that at the very least there seems to be nothing conceptually impossible about having norms that we think we must follow.

While it may be the case that a norm requires the capacity to "have done otherwise", this doesn't mean we have to think that we could have done otherwise.

Further, even if it is the case that we have to be able to have done otherwise to have a legitimate norm, false beliefs, motivated beliefs and biased belief are examples in which a belief *has* done otherwise; albeit outside the knowledge of the agent who has them.

Steglich-Petersen (2006) might object that this is only part of his concern. In comparing believing with promising, he notes that while it is "plausible to suppose that there are certain concepts that one cannot apply without accepting certain evaluations" it is hard to think that these "evaluations can necessarily motivate us to act in accordance with them" (Steglich-Petersen, 2006:506). In addition, he noted that in promising there was an intermediary step between making a promise and deciding to adhere to the inherent ought contained in promising.

In response to the argument in favour of a different type of norm, Steglich-Petersen (2006) might object that there are two things here to be worried about from the perspective of the Normative Account. The first, the one I've tried to deal with above, is that in other norm-following behaviour there is an inferential step in which we decide whether or not to follow the norm. The second is how the epistemic norm, particularly if we can't *choose* to engage with it, motivates us to act in accordance with it.

While a different type of epistemic norm might explain away the inferential step, it may well fail to account for the motivation of doxastic norms. For even if it is the case that we can't help but believe that *p* and believe that our belief that *p* adheres to the SoC for believing this doesn't explain the motivational power of the norm. Further, this constraint, if it exists, doesn't, in and of itself, explain why the SoC would always motivate us. Only that we have to think we adhere to the norm when we consider our second-order beliefs about one of our beliefs.

What reasons could we have to think that the norm is intrinsically motivating? In practical reasoning an agent can choose which reasons to follow and which not to.

Further, an agent can decide to ignore even very good reasons for a certain action in favour of other reasons. The same is not true in epistemic reasoning. While other factors undoubtedly play a role in determining what an agent takes as truth-bearing evidence, the fact of the matter is that in epistemic reasoning an agent is compelled to consider only reasons that he/she takes as being truth-directed. As mentioned above, he/she cannot ignore this evidence irrespective of his/her desire to do so. What this suggests is that when it comes to beliefs epistemic reasons are intrinsically motivating in a way that instrumental reasons are not when it comes to action.

Similarly, if we think that there's a case to separate epistemic norms from practical norms – and, further, think that these epistemic norms would have some similarity to epistemic reasons and reasoning – then perhaps it's the case that an epistemic norm is intrinsically motivating in the same way that epistemic reasons appear to be. If this is correct, then we have an answer to the first of Steglich-Petersen's (2006) first objection: epistemic norms are such that we can't help but think we've adhered to them, which mitigates the inferential step, and they are intrinsically motivating.[77]

However, this move is question begging. It assumes that epistemic norms are intrinsically motivating to show that they are intrinsically motivating. What needs to be explained is why epistemic reasons motivate us to act in accordance with them.

If an Aim Account is correct, then the constitutive aim of belief generates certain norms. These norms, in virtue of the fact that an Aim Account can postulate belief only aiming at Truth, would have to be followed in virtue of the constitutive aim of belief. As noted in the previous chapter:

> Postulating a goal for belief explains not only the authority of epistemic norms but also their motivational efficacy … the explanatory strategy behind it is perfectly sound: if believing were purposive, we could explain the above features of belief by reference to belief's purpose. (Owens, 2003:284-289)

[77] It might be objected that if the norm is intrinsically motivating then we should always expect the doxastic agent to be right. However, an agent can follow a norm and still fail to correctly achieve the desired-for state. If we're all brains in vats, for example, then it seems that many of our well thought-out and researched beliefs are nevertheless incorrect. Thus, an intrinsically motivating norm doesn't guarantee success.

Now, Steglich-Petersen's (2006) concern is that even if we can make sense of a constitutive norm in which there is no intermediary step between seeing the norm and deciding to fulfil the norm, this doesn't explain *why* that norm would motivate us. While an aim motivates us, or a purpose would drive something to fulfil that purpose, a norm doesn't appear to do that. Norms seem to be things that we decide to fulfil or not fulfil. In other words, something else would have to explain why we engage with the SoC for belief at all.

However, Steglich-Petersen's (2006) argument in this case appears to rest on the claim that norms cannot, in and of themselves, motivate us to act on them.

This line of reasoning seems similar to Hume's (1978), who argues that moral beliefs could not, in and of themselves, inherently motivate people to act (Hume, 1978:415). Only a belief coupled with a desire to fulfil the belief could cause someone to act on that belief (*ibid*.).

Steglich-Petersen's (2006) objection appears to follow similar lines, claiming that it would be hard to understand how a norm, in and of itself, could motivate beliefs to act in accordance with it. To suggest that this norm is intrinsically coupled with a desire for truth, suggests that what does the work is the desire and not the norm. As a result, we would need to abandon a Normative Account and instead talk about something like an Aim Account where what does the work is the desire's aim to reach the Truth.

Hume's (1978) claim is hugely contentious and presumably so is Steglich-Petersen's (2006). This is not a can worms that I have any desire to upend, nor do I have enough of a gut feel for one position over the other such that I would feel comfortable choosing one position in favour of the other.[78] As it stands, Steglich-Petersen (2006) certainly raises an objection a Normative Account would need to respond to.

That said, if a Normative Account is able to show that norms are intrinsically able to motivate beliefs to act in accordance with them, there seems to be no reason to think

[78] For an overview of the Humean Theory of Motivation see Connie Rosati's (2008) "Moral Motivation", in section 3.

that a Normative Account can't make sense of both the lack of an intermediary step as well as the motivational efficacy of epistemic norms.

In terms of the intermediary step it may be the case that we can't think of ourselves as taking practical reasons as a reason for belief even when we have in fact done so. What this suggests is that while we *can* do otherwise, we can't *think* we've done otherwise. Perhaps epistemic norms are constituted in such a way that we can't help but think we've adhered to them once we've formed a belief. In terms of the motivational force of an epistemic norm it is possible, though contentious, that the norm may in and of itself be intrinsically motivating such that doxastic norms have the motivational efficacy we encounter in beliefs.

If this is correct, then the Normative Account seems to be in a good position. What provides the explanatory tension needed in any account of the truth-belief relationship to explain the features of belief would be the constitutive norm that governs belief. Thus, what explains, for example, why only epistemic reasons count as reasons for belief is the fact that in order to meet the constitutive norm of belief the only way to do so would be to take only epistemic reasons into account.

It escapes Shah's (2003) fatal dilemma for Purpose Accounts by highlighting that a belief ought to meet the SoC, but can get it wrong: an explanation that's not available to Purpose Accounts. In doing so, it can make sense of motivated beliefs by admitting that we can fail the norm of correctness, but nevertheless keep transparency as this would still be the only way of trying to adhere to the norm.

However, rather than put this straight into the problems faced by Aim Accounts, it avoids those by postulating an explanatory tension with Truth in virtue of norms, rather than have the belief itself sit in tension with Truth. What this does is to avoid placing belief in an odd position where the direct link between belief and truth is compromised.

Thus, if Normative Accounts are able to address the concerns raised by Steglich-Petersen (2006) above, then they seem well placed to account for several features of

belief. What I explore in the following section is the ability of Normative Accounts to deal with the other features of belief I have discussed in Chapter 2.

6. Accounting for the "other" features of belief

Normative Accounts gain traction in virtue of the fact that they are able to explain both transparency and motivated beliefs in virtue of beliefs following a prescriptive and constitutive norm of correctness.

In this section I examine what happens if we take the Normative Account seriously. I argue that taking this model of belief seriously leads to two failures. The first is that while there are some cases in which false beliefs are easily explained there is a deeper question about how this account makes sense of any belief based on inconclusive evidence. The second is that while a Normative Account may be able to explain three of the four instances of *idées fixe* I examine in Chapter 2 it cannot make sense of beliefs that maintain in the face of unanswerable yet unbelievable arguments and so should be abandoned as the correct account of the truth-belief relationship.

6.1. False beliefs and beliefs based on inconclusive evidence

In accounting for motivated beliefs, Normative Accounts point to the fact that while the SoC for belief governs beliefs and, therefore, beliefs ought to adhere to it, this describes "how beliefs ought to be, not how they actually are" (Steglich-Petersen, 2006:506). This move equally seems to explain false beliefs: a false belief, like a motivated belief is simply one that has failed to meet the constitutive norm of correctness.

But how has this occurred? In the easy-to-describe cases what has happened is that despite the best efforts of the agent or the belief he/she/it has been misled. Consider the following case. I am trying to find out when the next Liverpool game is. I check online, cross-checking the information not only on my local satellite provider's website, but with various other football sites. All of them agree that the game will be screened at 19:30 (CAT) on the upcoming Saturday. However, unbeknownst to me, and to the various sites I've checked, there was a clerical error at the Premier League

offices and the names Liverpool and Stoke were accidentally switched round on the official TV schedule that the Premier League released. This error, coming as it does from the official source, has been dutifully copied across to all other sources. Because of this error all sources of possible information at my disposal agree, incorrectly, that the next Liverpool game is at 19:30 (CAT) on Saturday.

In this case, through no fault of my own – and no fault of the belief trying to meet the constitutive norm of belief – I have formed a false belief: one that despite doing more-or-less everything I could do to ensure that the belief was True is nevertheless false. In this case, it seems that the belief's inherent SoC would be satisfied and a belief would be formed. All the epistemic norms that go along with forming a belief as a result of the constitutive norm of belief appear to have been met and so it comes as no surprise that a belief has been formed even if, in this case, the belief fails to meet the SoC.

But what about those cases where my false belief is based not on, *ceteris paribus*, conclusive evidence, but instead on inconclusive evidence? Going back to the time of the Liverpool game, let's consider a case in which I'm musing aloud in a shopping centre about when the Liverpool game is. A stranger, overhearing my musing, tells me it's at 19:30 on Saturday. As a result, I form a belief.

This isn't a particularly unusual situation, as Owens (2002) notes, "people form beliefs on the basis of inconclusive evidence all the time …" (Owens, 2002:393). But, if a Normative Account is correct, how is it possible to form beliefs on the basis of inconclusive evidence? A Normative Account claims that a belief is formed in virtue of an adherence to a norm that a belief is correct if and only if it is True. The belief, itself, governed by this norm, can only be formed once it feels that the SoC has been met.

While the Normative Account doesn't demand that True beliefs are formed in the "right" way, only that the resultant belief meets the SoC for belief, given that the SoC operates prescriptively how do we make sense of beliefs formed on inconclusive evidence? It seems that to meet the SoC one would be required to engage in practices that try to meet the SoC – something that appears not to happen when we believe on

inconclusive evidence. That a belief based on inconclusive evidence may happen to meet the SoC in some cases doesn't undermine the concern that forming a belief on inconclusive evidence would seemingly be ruled out by a Normative Account in virtue of the controlling norm of correctness.

This norm, it seems reasonable to suggest, in virtue of what it is, would require a relatively high degree of evidence in order to ensure that the norm is met. Failure to engage in practices that aim quite rigorously at Truth would be to fail to adhere to the SoC. But this is not what happens. I can and do form beliefs on very little evidence.

One way to cash out this concern is that we appear to be in danger of falling back into something like the dilemma Shah (2003) notes is fatal to Purpose Accounts. Either the regulatory force of the doxastic norms must be strong enough to indicate that the prescriptive norm that the doxastic norms are trying to discharge is the SoC – in which case it appears unable to explain beliefs based on inconclusive evidence – or the prescriptive force must be weak enough to explain beliefs based on inconclusive evidence – which suggests that the norm that beliefs are trying to discharge is not the SoC.

One way out of this for Normative Accounts is to claim that while many of the doxastic norms are Truth-directed, they don't guarantee Truth. So, while following doxastic norms means we are more likely to reach a True belief than if we failed to follow the norms, these norms, while aiding a belief in fulfilling the norm of correctness, don't guarantee that result. Transparency, for example, merely requires that the reasons taken as reasons for the belief are seen as epistemic ones. But this doesn't guarantee that the reasons *are* epistemic reasons, or, even if they are epistemic reasons that they're the *right* epistemic reasons for the belief in question.

If this response to the above objection is correct, then Normative Accounts must claim that the prescriptive force of the SoC is such that our doxastic norms operate at a level that lends itself to, but doesn't guarantee, satisfying the norm of correctness.

However, I can't see that this solution works. The Normative Account states that in forming a belief, this process is governed by the constitutive SoC for belief. In

forming a belief, therefore, I am constrained to believe that *p iff p* is True. But given what the SoC is, it would seem that my belief would *push*, as it were, for far greater epistemic rigor such that believing based on inconclusive evidence oughtn't to be possible. The norm that beliefs are trying to discharge is not 'to believe that *p* if in some possible world *p* is correct'. Instead it's 'to believe that *p* if and only if *p* is in fact True'.

It seems that given this SoC every instance of, say, perceptual believing would, as a result of the constraints demanded by the SoC, end up being a case of deliberative believing. To believe that there is a glass on the table based merely on the perception that there is a glass on the table, would seem to put the belief in danger of failing the prescriptive and constitutive norm of belief. Given that this norm governs the very processes that make up a belief, shouldn't we expect under this account that I would feel compelled to investigate this claim, any claim, any candidate belief, with far more rigor than I see myself actually doing?

If a belief is governed by the prescriptive condition to believe that *p iff p* is True, then it seems that I couldn't believe that Liverpool are playing on Saturday at 19:30 because someone I don't know told me in passing that this was the case. To do so would be to believe almost in the face of the SoC for belief. But we do believe on inconclusive evidence and this undermines the claim that our beliefs are regulated in virtue of a constitutive norm of correctness.

How would a Normative Account explain this? One option would be to deny that the norm operates prescriptively. If the norm is something that indicates success after the belief has been formed, then there is no reason to think that this norm would operate in such a way as to try to ensure that the belief actually meets the norm. What is causing problems for the Normative Account here is the fact that if the constitutive norm of correctness for belief operates prescriptively, then it is hard to see how we're able to form beliefs on inconclusive evidence.

If we deny that the norm operates prescriptively, then the worry is solved and what we have is a way of adjudicating whether beliefs are correct or incorrect without thinking that this SoC would have any other constraints on the processes of belief.

However, this move is surely fatal to a Normative Account. If the norm is descriptive, rather than prescriptive, then it couldn't explain why we are compelled to follow the norm. What does the work in a Normative Account is the inherently motivating force of a constitutive norm operating prescriptively. It is the prescriptive force of the norm that explains features of belief like transparency. In this case the Normative Account must decide whether to give up being able to follow any of the epistemic norms as they pertain to belief, or must give up the ability to explain beliefs formed on inconclusive evidence: neither of which looks particularly appealing.

Another route is to appeal to the first-person-third-person divide. When I form a belief I take that proposition as being True and in doing so, from my own perspective, the evidence I have for my belief is conclusive. Thus, when I form these beliefs, despite the paucity of evidence from a third-person perspective, I nevertheless take the proposition in question as being True. As I noted in Chapter 2, whatever the correct account of the truth-belief relationship is, it is the case that whenever I form a belief I do so in virtue of the fact that I take the believed proposition to be True. As far as the believer is concerned, therefore, the norm has been satisfied. Normative Accounts might claim, therefore, that since the believer takes the norm to be satisfied (consciously or unconsciously) the norm is satisfied. This explains how we can make sense both of a prescriptive and constitutive norm of belief that regulates our beliefs as well as beliefs that are formed on inconclusive evidence.

However, this move seems to beg the question of how the belief was formed in the first place. What's at issue here is how the believer might think that *p* is True if the evidence for *p* is inconclusive. Stipulating that when we form a belief we do so in virtue of *taking p* as being True doesn't answer the question how it is possible for us to take something as being True given a) that the evidence for *p is* inconclusive, and b) there is a constitutive norm of correctness governing our beliefs prescriptively to meet this norm.

In virtue of what does the agent or belief take it as being True? If the SoC demands nothing less than the Truth, it seems hard to understand how we could form beliefs

based on inconclusive evidence. Stating that I do begs the question of how I form these beliefs in the first place.

6.2. Normative Accounts and *idées fixe*

In Chapter 2 I examined four different types of *idées fixe*.

1. Long-lasting bigoted beliefs.
2. Delusions
3. Beliefs that maintain in the face of unanswerable and yet unbelievable arguments

And

4. Things "I was going to believe anyway"

All of these are cases, I argued, in which the agent holds onto a belief in the face of, sometimes overwhelming, counter evidence.

How would a Normative Account explain this evidence-resistant feature of belief?

In the case of options 1, 2 and 4, the agent feels that belief meets the relevant epistemic norms. The bigot not only takes his/her bigoted belief as being True, he/she is able to provide an epistemically rationalising story, provide (what he/she takes as) epistemic reasons for his/her belief and so on. The same is true of someone who suffers from certain delusions. Similarly, in the case of beliefs "I was going to have anyway" these tend to be things that I "feel" are True even if I'm unable to always provide the type of defence of that belief that would satisfy an epistemic ethics committee. If this defence is sufficient to get Normative Accounts out of the trouble posed by beliefs formed on inconclusive evidence, then perhaps this is the way to explain these *idées fixe* as well.

As I've argued above, this move doesn't work. It seems to be inherently question begging. Therefore, it seems to me that these three types of *idée fixe* will stand or fall with the defence of beliefs based on inconclusive evidence. I've argued that the defence above fails and so not only do Normative Accounts fail to make sense of

beliefs based on inconclusive evidence, they also fail to make sense of at least these three types of *idée fixe*. How could a Normative Account with a SoC operating prescriptively make sense in any way of beliefs that stand in the face of what appears to be conclusive evidence to the contrary? They can't. If, however, I'm wrong and Normative Accounts can make sense of beliefs based on inconclusive evidence in virtue of an agent taking it to be True, then a similarly perhaps the same strategy will serve to explain how to make sense of these *idées fixe*.

As a result, I turn to the case described by 3. This type of *idée fixe*, I think, represents an explanatory challenge that a Normative Account cannot meet irrespective of whether the explanation of beliefs based on inconclusive evidence goes though.

What could potentially save a Normative Account from the concern that we form beliefs based on inconclusive evidence, as well as explain the *idées fixe* detailed in 1, 2 and 4, is the fact that the evidence for these beliefs satisfies the believer. While this is somewhat dodging the issue, since it immediately raises the question of why it satisfies the believer, let's assume that something like this strategy will work for a Normative Account. What makes the *idée fixe* detailed in 3 above so interesting is that the agent who believes that *p* may feel that the evidence that he/she has for *p* is not conclusive.

Brian Ribeiro (2002), discussing his belief in an external world in the face of external world scepticism, believes that an external world exists, but thinks that he has no epistemic reasons for this belief at all. The unlikely, but potential defence of Normative Accounts to explain beliefs formed on inconclusive evidence is not available in this instance. In this case we have a belief that by the agent's own lights has failed the norm of correctness and yet still maintains.

In these cases, given a prescriptive norm to believe that *p iff p* is correct, which operates below the intentions of the agent, at the very least we should expect to see the agent's belief suspend in the face of an unanswerable yet unbelievable argument. That they don't, suggests that whatever process governs our beliefs it is not the norm of correctness operating in the way that a Normative Account is committed to.

A Normative Account claims that the SoC for belief operates prescriptively such that in order to believe that *p* one's belief that *p* must have met the SoC. But in the case of beliefs that maintain in the face of unanswerable and yet unbelievable arguments this norm seems poorly placed to explain the existence of these beliefs.

7. Conclusion

A Normative Account of the truth-belief relationship attempts to explain the features of belief in virtue of a constitutive and prescriptive norm of correctness.

Often presented as an account of deliberation, in this Chapter I have tried to clear up any confusion by reiterating the distinction made in Chapter 2 between an account of the truth-belief relationship and the projects a doxastic agent can engage in. With this distinction in place I've laid out a Normative Account of the truth-belief relationship.

I then examined several potential concerns. The first thing I noted was that this account is able to escape both some of the problems of the Purpose Account as well as some of the problems of Aim Accounts, suggesting that this account may well be a good explanation of the truth-belief relationship.

The second set of objections I examined comes from Steglich-Petersen (2006). I argued that of the three objections he puts forward, only the first one is an objection to a Normative Account of the truth-belief relationship, while the others, as they stand, appear to be objections to a normative understanding of deliberation.

Steglich-Petersen's (2006) first objection argues that a Normative Account must, in virtue of postulating a norm that is followed, give up the ability to explain not only transparency, but the motivational force of doxastic norms. I argued, using a clue found in Gibbard's (2005) work, that we can possibly get round the first of Steglich-Petersen's (2006) objections if doxastic norms, like doxastic reasons, are fundamentally different to their instrumental counterparts.

I argued that this move may work, but that the objection would be something that proponents of Normative Accounts would need to address to save the account. I then

examined whether a Normative Account can make sense of several other features of belief.

I began by examining this account's ability to make sense of false beliefs. I argued that while there are some cases in which Normative Accounts are able to explain these beliefs the account seems hard pressed to explain beliefs (True or false) based on inconclusive evidence.

However, while there may be ways out of these two objections, I argued that a Normative Account cannot explain beliefs that are held in the face of unanswerable and yet unbelievable arguments.

This failure of Normative Accounts to capture several features of belief I believe means that these accounts, as with the teleological accounts we examined in the previous chapter, should be abandoned.

In the following chapter I put forward my account of the truth-belief relationship.

Chapter 5: Aiming at rationality

1. Introduction

In Chapter 2, I argued that any account of the truth-belief relationship must, in virtue of that relationship, be able to explain the features of belief. Should a theory of this relationship fail to explain one of the features of belief, it would undermine the account in question.

In Chapters 3 and 4 I examined two models of the truth-belief relationship found in the extant literature: teleological models, which I explored in Chapter 3, and normative models, which I explored in Chapter 4.

Both of these models argue that the features of belief are explained in virtue of an explanatory tension between beliefs themselves and Truth. In the case of teleological accounts this tension is either that beliefs aim at the Truth, or that beliefs have the purpose of Truth. In Normative Accounts, the claim is that beliefs follow the prescriptive norm of Truth.

In all of these cases I argued that these accounts fail to make sense of various features of belief laid out in Chapter 2. I argued that there are two main failings with these accounts. The first is that by trying to explain the features of belief via a tension between beliefs and Truth we lose what it's like to be a belief. The second is that none of these accounts appear to be able to explain evidence-resistant beliefs (*idées fixe*). These two failures undermine these accounts as the right way of understanding the truth-belief relationship.

The failure of the two existing models of the truth-belief relationship, therefore, requires the development of a third. In what follows, I propose an alternative account of how to understand the truth-belief relationship. This model, I feel, best captures the doxastic experience by being able to account for the features of belief I argued for in Chapter 2.

My model is made up of two independent, mutually-compatible parts. The first part denies that beliefs and Truth are in tension with one another, while the second part suggests that beliefs are part of a doxastic system that has the goal of Truth. While this may seem unnecessarily complicated what I hope to show below is that the explanatory power of this model makes this a more appealing model of the truth-belief relationship than the other two models we've examined.

I've argued in Chapters 3 and 4 that the explanatory tension that other accounts place belief and Truth into ultimately undermines these accounts of the truth-belief relationship. My solution, therefore, is to dissolve the tension between beliefs and Truth (the Tensionless Account). However, while this Tensionless Account has many benefits, as I argue later on, it fails to explain how it is that beliefs follow doxastic norms.

The second part of my account, therefore, suggests that the explanatory tension is provided by a relationship a doxastic *system* has with Truth. The theories we've looked at in the previous chapters detail the relationship between a belief *itself* and Truth. My approach argues for a tensionless understanding of the link between belief and Truth, within the broader context of a system that controls our doxastic lives.

The suggestion goes that this doxastic system, whose job it is to form accurate beliefs, is just one part of our mental landscape. I suggest that this doxastic system, in virtue of its links with an agent's mental landscape as a whole, is shaped by the agent's other, not necessarily epistemic, commitments. This doxastic system, I argue, balances the inherent purpose of Truth against these other commitments. I argue that as a result the doxastic system aims to rationally achieve the target of Truth, but that this target is tempered by other cognitions, emotional states and conations that make up a rational agent. As a result, the system, with the goal of Truth, while maintaining a certain level of internal coherence, can be said, *de facto*, to aim at rationality, rather than Truth.

I argue that a Tensionless Account of belief, understood as part of a doxastic system that aims at rationality in this rather convoluted way is best able to make sense of both

the Truth-tracking features of beliefs as well as account for the truth-adverse features of belief that might puzzle us.

In Section 2 I explain and argue for a Tensionless Account. In Section 3 I lay out what I mean by the doxastic system and provide two arguments in favour of this system. In Section 4 I argue that a doxastic system with the inherent goal of Truth would be able to provide the explanatory tension needed to make sense of several features of belief, while avoiding the concerns of Purpose Accounts laid out in Chapter 3. In Section 5 I argue that a doxastic system is shaped by the other commitments an agent has. In Section 6 I critique my approach. I suggest that while my account is able to account for more of the features of belief than other accounts, there are holes in the theory. In particular I'm not sure how my account makes sense of holding onto beliefs that don't fit into the rational schema I argue for in Section 5.

2. Beliefs, Truth and the Tensionless Account

The other accounts of the truth-belief relationship we have looked at place belief and Truth in a specific relationship with one another in order to be able to explain various features of belief.

What does the work in these accounts is the tension between Truth and belief such that the gap between the actual Truth and what an agent takes as Truth (his/her belief) is the explanans for whatever doxastic explanandum we're looking at.

A Normative Account, for example, places belief in a norm-following relationship with Truth in order to make sense of, in the case of Nishi Shah and David Velleman (2005), transparency. The tension between the belief and the norm generated out of Truth is what, this account claims, explains transparency.

However, as I've argued in the previous chapters, by placing beliefs and Truth in these tension-creating relationships we lose the ability to explain important features of belief, and as a result what it means to call a mental state a belief.

In this section I suggest that the way to understand the truth-belief relationship is as a tensionless one. The claim is that the concept of belief (which is what a Tensionless Account is) uncomplicatedly explains and helps us understand several features of belief. While this might be the moment where one exclaims "obviously the concept of belief provides the features of belief!", this section highlights that other accounts essentially *lose* the concept of belief the moment they postulate an explanatory tension between beliefs and Truth. A Tensionless Account in virtue of being able to explain the features of belief, therefore, actually provides reasons for accepting this as the proper account of the relationship between belief and Truth. Let's turn now to the deficiencies of other accounts of the truth-belief relationship.

2.1. Aims and the gap to truth

One of the biggest problems with the Aim Account is that it seems to place Truth in the wrong relation to belief.

In order to provide the explanatory tension that explains how beliefs regulate for truth, for example, Aim Accounts must claim that we form a belief with the aim or intention of *p* being True. Thus, what explains our beliefs regulating for truth is that we form a belief that by aiming at Truth is able to regulate for truth.

However, this doesn't capture what it is to believe something. When we believe something we *already* think it's True. We don't form a belief with the hope that at some point in the future it will be True, we form a belief precisely because we already think it is True. The explanatory gap, therefore, fails to capture what it means to be a belief.

2.2. Purpose: infallibility or the explanatory gap

Under a Purpose Account, Truth would have to pull, as it were, the propositional attitude towards it. What would, in theory, explain that beliefs regulate for truth would be the mental state being pulled into correctness.

But, if the purpose of belief is Truth, then it seems that the only way to believe anything at all is if it is actually True. This suggests that our beliefs, in virtue of being beliefs, must be infallible. However, given that this is absurd, how do we make sense of beliefs having the purpose of Truth?

The suggestion would be to resort to the explanatory gap that Aim Accounts appeal to; namely the distinction between what we believe (the truth) and the actual Truth. However, to appeal to this gap is to fall into the same trap as Aim Accounts above. To believe something, under this account, would be to think that one has achieved the purpose of belief. But if the very belief thinks that the purpose has been achieved how does the inherent purpose provide any explanatory tension? Under this account the beliefs, when formed, in some way must continue to think that the purpose of belief has not been discharged, which leads to a case in which we form beliefs with the hope they will be True at some stage. But since to believe that *p* is to take *p* as being True, this gap is incommensurate with what it means to be a belief.

2.3. The failure of norms

Under a Normative Account, what guides the belief processes is the prescriptive force of the standard of correctness for belief. What a Normative Account claims is that this standard of correctness is a norm that operates prescriptively to guide the processes of belief.

There are several problems with this account.

The first is that a Normative Account has to explain how it is that we have an intrinsically motivating norm. As Asbjørn Steglich-Petersen (2006) argues, our experience with norms indicates that they are things that we can decide to follow or not. But doxastic norms are not like this and a norm-following account seems poorly placed to explain this.

The second is that if what constrains a belief is the norm of correctness, it is hard to see how it is ever possible to form a belief on inconclusive evidence. A belief is constrained to meet a norm that demands Truth. If our belief processes are governed

to form a belief only if it is correct, then we shouldn't be able to form beliefs on inconclusive evidence. But we form beliefs on inconclusive evidence all the time, thereby undermining a Normative Account as the right way to understand the truth-belief relationship.

The third problem is that Normative Accounts fail to make sense of *idées fixe*.

All three failures undermine the Normative Account as the right way to understand the truth-belief relationship.

The core problem in all of these accounts, I'm suggesting, is that they place beliefs themselves in an explanatory tension with Truth. My suggestion, therefore, is to dissolve this tension in order to save what it means to be a belief.

2.4. Truth, belief and a lack of explanatory tension

To see how this works, let's go back to the concept of belief. To believe that *p* is to take *p* as being True for no extrinsic reason that's transparent to the agent, but simply because one takes *p* as actually being True.

In trying to work out when the TV show *The Amazing Race* is screened, while the investigative action is done to meet your desire to find out when the show screens, your belief (assuming you form a belief) isn't formed with some extrinsic end in mind. It's formed solely because your investigation has led you to take it as True that *The Amazing Race* screens at a specific time. When you believe that *The Amazing Race* is on television tonight, you don't believe it hoping it *will* be True. Instead, you believe it, because it is something that you take as *being* True.

As I've mentioned a few times, what does the explanatory work in other accounts is that a belief is placed in tension with Truth. In the case of Aim Accounts beliefs aim at Truth, in the case of Purpose Accounts beliefs have Truth as their purpose and in the case of Normative Accounts, beliefs follow the Truth norm. As I've summarised above, it's in virtue of this tension between beliefs and Truth that these accounts get themselves into trouble.

My suggestion is to dissolve the tension in the relationship between belief and Truth. The only relationship that beliefs have with Truth, I claim, is the conceptual one. That when one believes that *p* one takes *p* as True.

What advantages does this theory of the belief-truth relationship have?

The argument is very simple. By placing belief and Truth in tension with one another other theories fail to capture what a belief is and this must be seen as a failure of this type of move. The solution, therefore, is to negate this tension between belief and Truth. To put the two into a non-dynamic relationship with each other such that what one thinks is True is a belief and a belief is what one thinks is True.[79]

What makes this move less than obvious is that other accounts fail to capture the concept of belief by postulating a further tension. The obvious advantage of the Tensionless Account is that it fits in with what a belief is: when you believe that *The Amazing Race* is on television tonight you do so because you take it as actually being True that this state of affairs is correct. Not in the future, as other accounts seem to commit us to, but True, *simpliciter*.

But while saving the concept of belief is clearly an important part of this account, it is not the only thing that an account of the truth-belief relationship needs to do. In addition, the truth-belief relationship must be able to explain the features of belief.

The features of belief I laid out in Chapter 2 were that:

1. The relationship between belief and truth is direct. (The first way one could make sense of transparency).

[79] As I spelt out in Section 2 of Chapter 2 (on what a belief is), this is not the same as claiming that the Truth just is what we believe. These anti-realist theories of Truth are making a claim about Truth. I, on the other hand am making a claim about what belief is and its relationship to Truth. I do not think that believing *p* in anyway confers Truth on the world, nor do I think that believing that *p* guarantees that *p* is actually True (though, the agent must think that it is). The claim I'm making is that the propositional attitude we have when one takes *p* to be True is called a belief and that that is as far as the relationship between beliefs and Truth goes.

2. The scope of belief is vast and varied. That agent A can believe that *p* and that agent B can believe that not-*p* and that both *p* and not-*p* count as beliefs.
3. Only (what an agent takes as) an epistemic reason counts as a reason for belief (the second way one could cash out transparency).[80]
    a. As a corollary, unless we're in the face of something like Brian Ribeiro's (2002) unanswerable yet unbelievable arguments, when you have a belief that *p* you find that you tell an epistemically rationalising story about that belief.
4. Epistemic reasons have a direct link with belief, such that if you find epistemic reason *r* for *p* compelling, *r* would directly cause you to believe that *p*, without the need to weigh *r* against instrumental reasons for or against *p* (the third way of understanding transparency).
5. Beliefs are involuntary.
6. Beliefs regulate for truth.
7. We sometimes, unbeknownst to us, form beliefs that are based on non-epistemic reasons, i.e. motivated beliefs.
8. Some of our beliefs are biased.
9. We form false beliefs.
10. We form beliefs on inconclusive evidence.

And, finally,

11. In some instances our beliefs maintain in the face of counter evidence (*idées fixe*).

A Tensionless Account explains why the link between beliefs and truth is direct. Given what it is to be a belief, it's obvious why "[w]hen you change your mind about what is true, you have therein, *ipso facto*, changed your beliefs" (Hieronymi, 2006:53). If the relationship between belief and what we think is True is identical, then, axiomatically, when one changes one's mind about what is True one simultaneously changes one's belief. While this cashing out of the truth-belief relationship doesn't explain *how* we change our minds about what is True (more on this later), it does explain why our belief automatically changes when what we think is True changes.

[80] When referring to the reasons that an agent has for a belief I use the term "epistemic reason" to mean "what an agent takes as an epistemic reason".

In addition to being able to account for the direct relationship between beliefs and truth, this account of the truth-belief relationship also makes sense of the scope of beliefs. Given that to believe that *p* is to take *p* as being True, the scope of beliefs is limited to whatever someone can take as being True.[81]

Transparency, as argued for in Chapter 2, can be understood in three distinct ways. As the direct link between beliefs and truth; as the notion that only epistemic reasons count as reasons for belief; or that epistemic reasons have a direct impact on our beliefs.

The first of these possible understandings of transparency has already been accounted for above. The question then becomes, does a Tensionless Account make sense of the latter two formulations? Well, if what you believe is identical to what you take as True, then in asking the question whether to believe that *p* what you are actually asking is whether *p* is True or simply whether *p*. What makes the doxastic question transparent to the factual question is that there is no space between what one takes as True and what one believes. The questions are the same thing.

This explains why only epistemic reasons would count as reasons at all. In order to have a belief we would have to have a mental proposition that one thought of as True for no extrinsic reason that was transparent to one, but simply because one took the proposition as True. While one could potentially be deceived or misled into taking a non-epistemic reason as an epistemic reason, from your own perspective it would nevertheless be the case that you could not take anything other than epistemic reasons into account when forming a belief. If you did, it would be hard to see that you could think of your belief as a belief. However, while a Tensionless Account explains why

[81] While I've not explicated this point previously it seems to me that if Purpose Accounts and Normative Accounts decide to avoid the worrying gap I've highlighted above in favour of something like infallibility we could not explain why different people, legitimately, have different beliefs. Assuming something like a realist perspective on Truth (which I am), if we all had infallible beliefs I would expect to see very little differences in our beliefs. That we have different beliefs, or could easily have different beliefs, suggests that this is certainly not the way to understand either Purpose or Normative Accounts. A Tensionless Account, by acknowledging that what we believe is what we *take* as being True, explains this feature comfortably.

we can only have epistemic reasons as reasons for a belief, it doesn't explain what stops us from ignoring these reasons.

The third feature of transparency is that epistemic reasons for the truth of some claim directly impact on what we believe such that we don't need to consider instrumental reasons as well as epistemic reasons. A person with their full mental faculties intact who is suddenly confronted with incontrovertible evidence that their house is on fire forms a belief directly as a result of this evidence. While a Tensionless Account explains why only epistemic reasons for *p* count as reasons to believe that *p*, it cannot explain how these reasons motivate us to act (believe) in accordance with them, nor why they would motivate us directly. I argue in Section 3 that what explains these features of belief is a doxastic system.

Moving on, a Tensionless Account can explain why one needs to *think* that one has an epistemic reason for a belief. This is a separate point to the claim that only epistemic reasons count as a reason to believe. The latter claim explains what reasons you can take as a reason to believe, the former is what you can say about one of your beliefs. If to believe that *p* is to take *p* as being True for no extrinsic reason that is transparent to you, then in telling a story about this belief it seems you must tell an epistemically rationalising story.[82]

In addition, a Tensionless Account can explain why, if not how, beliefs regulate for truth. If you stop thinking something is True, then, given the relationship with belief, you stop believing it.

[82] I argued in Chapter 2 that there are unusual cases where one's belief runs up against an argument that one both sincerely finds unanswerable and also finds unbelievable. In these cases we might be forced to conclude that our belief is held for non-epistemic reasons. Thus, the truth-belief relationship needs to do two things. One, it needs to account for the fact that, *ceteris paribus*, when you examine a belief you hold you tells an epistemically rationalising story about that belief. And two, it needs to explain how it is that in the face of unanswerable and yet unbelievable arguments you may find you have a belief that you yourself think you don't have an epistemic reason for. It is worth noting that in these latter cases you don't provide a non-epistemic reason for your belief. You don't, for example, say of your belief in the external world "well I believe that the external world exists because I like ice-cream". (Well not with any degree of doxastic seriousness at any rate). Instead, in this situation you are left baffled by why you believe that *p* at all. You acknowledge that you have no epistemic reasons for *p*, and so your belief must be based on a non-epistemic reason, but is seems you would be hard pressed to provide a non-epistemic reason for the belief; even if you acknowledge there must be one. I believe that a Tensionless Account is able to explain the former phenomena. What this account can't do, as of yet, is explain the latter phenomenon. I explore a way to answer this problem below.

A Tensionless Account is also able to explain some forms of wishful thinking and simply getting it wrong. If an agent takes something as True, he/she believes it. If, for whatever reason, the agent's belief is based on non-epistemic reasons (opaque to the believer), then the fact remains that the agent takes it as True and so believes it. Similarly, while the belief may be objectively false, if the agent takes it as True, he/she believes it.[83]

Similarly this account can explain why, though not how, we form beliefs on inconclusive evidence. I take *p* as being True and thereby believe it. While *p* may be underdetermined by the evidence, because I take it as True I believe it.

There is also a very trivial sense of how a Tensionless Account is able to account for the *idées fixe* I've argued for in Chapter 2. In all cases of *idées fixe* the agent takes the believed proposition as True. A Tensionless Account, therefore, is able to account for these *idées fixe* as genuine beliefs. As we have seen in Chapter 2, Velleman (2000) tries to get round the problem of *idées fixe* by simply denying that they are beliefs at all. A Tensionless Account avoids having to argue for that move. Of course it doesn't and, in fact, cannot explain how they exist, but it can make sense of the fact that these are beliefs. Something that the other accounts we've looked at can't do.

What makes a Tensionless Account attractive, therefore, is that it is able to make sense of a number of features of belief. By placing truth in a non-dynamic relationship with belief, this account is able to account for several of the features of belief without necessitating the type of gap that is created in the other theories of the truth-belief relationship. In doing so, the account provides reasons for why we should adopt this as part of the correct account of the truth-belief relationship.

[83] To give a trite example: consider playing Scrabble. You have a combination of letters that fails to constitute a real word. You want to put down a high-scoring word and so convince yourself that you've seen this word before and believe that this word "exists". In this case your thinking has become "hostage to your desire and wishes ..." (Williams, 2002:83) and you've incorrectly formed a belief. Under a Tensionless Account this is easy to explain. Because the relationship between truth and belief in this case just is what a belief is, then what accounts for the belief formation is that the agent has taken the proposition believed as being True, for whatever reason, despite it not actually being True. Of course, it does the raise the question of how we mistake non-epistemic reasons for epistemic reasons, again, something I explore below.

However, while a Tensionless Account explains a number of features of belief, some sort of tension is required to explain other features of belief. In virtue of what, for example, do beliefs regulate for truth? Why is it that we can't ignore epistemic reasons? Explaining that because to believe that *p* is to take *p* as being True explains why it's the case that when you stop thinking *p* is True you thereby stop believing it. But it doesn't explain *how* you stop thinking *p* is True. Nor does it explain why we can't ignore epistemic reasons – even if you want to. Some sort of explanatory tension is required to do this.

Similarly, while a Tensionless Account doesn't rule out *idées fixe* in the way that other accounts do, how these beliefs exist in the face of counter evidence still needs to be explained.

To address these concerns I turn to the second part of my account, which suggests that we should see beliefs as part of a doxastic system. I argue that it is this system that provides the tension needed to explain why doxastic norms have the motivational power they do. Further, I argue that this system is better able to make sense of our doxastic experience than the existing models of the truth-belief relationship.

3. Doxastic systems

By suggesting that there is no tension between beliefs and Truth I believe we place belief and Truth in the right relationship with one another. However, in doing so the explanatory tension that explains how our beliefs behave the way they do is lost. My suggestion is that what provides this explanatory tension is the doxastic system, aiming to rationally achieve the goal of a tempered Truth.

In this section I lay out what I mean by a doxastic system before I present two arguments in favour of this system. The first argument examines a claim made by Noam Chomsky (1959) about language acquisition (Chomsky in Swartz, De La Rey & Duncan, 2004:230). Chomsky argues that the way that children learn languages suggests a pre-existing structure in the mind/brain that aids the learning of language. I argue that something similar to Chomsky's language-acquisition structure exists with

regards to beliefs – suggesting that something like the doxastic system could be part of an evolutionary story of human cognition.

That said, it is important to note that my argument for a doxastic system does not rest on this evolutionary claim. My second argument is that a doxastic system *qua* a doxastic system is better able to describe the features of belief, and, therefore, provides a better account of the truth-belief relationship than the other accounts we have examined.

I begin by sketching out what I think a doxastic system might look like. Having done so I provide an argument for how this system could be understood as an evolutionary system. I then argue that a doxastic system that governs the processes of our beliefs makes more sense than the alternatives suggested in the literature. In Section 4 I show how this system might provide the required explanatory tension, before arguing in Section 5 that this system should be seen as aiming at rationality.

3.1. Making sense of a doxastic system

A doxastic system, I suggest, is a system that helps us order and make sense of the world epistemically. This doxastic system is the system that is responsible for the formation, revision and retention of our beliefs: telling us what and who to believe, what to deduce or infer, and so on. This system governs our doxastic lives in much the same way that the homeostatic system regulates our body temperature.

A rational being, the proposal goes, interacting with the world would need to know how the world is. A system that differentiated between truth-bearing cognitions and non-epistemic cognitions and conative states would help us form, revise and retain those cognitions that helped us make sense of the world epistemically – determining what is True, or appears to be True.

In addition to beliefs, my suggestion is that what's needed to make sense of the world epistemically are all cognitions, conations and possibly epistemically relevant feelings (intuitions, hunches, feelings) that help us determine what is and what isn't True. Unlike the theories we've examined, which appear to require beliefs to be the sole

point of call in determining how the world is, the suggestion here is that beliefs are just one part of a system that governs our epistemic outlook.

So what would make up a doxastic system? Without claiming that this is an exhaustive list, my thoughts are that this system would include: our perceptual faculties; our deductive, inductive and abductive reasoning; hunches and intuitions; desire to get things right and, of course, our beliefs. In short, everything that is involved in forming, revising and retaining a belief including other beliefs; whether they are background beliefs or beliefs that function as premises in an argument structure.

It is this system, I'm suggesting, that allows us to: differentiate between epistemic reasons and non-epistemic reasons; correctly incorporate and interpret relevant perceptual stimuli; fit new beliefs in with existing beliefs; follow arguments; deduce/abduce/infer certain claims from other claims or evidence; take on or reject testimony; and so on. In doing so the system links relevant cognitions, perceptions, background beliefs, epistemic dispositions, arguments and other potentially Truth-directed cognitions, conations, feelings, structures, mental states and dispositions as part of a process that forms, retains or revises a particular belief. It is this system that explains the motivation of epistemic reasons, and which explains why our beliefs follow epistemic norms more generally.

3.2. Doxastic system and evolution

One of the reasons that make the postulation of a doxastic system, or sub-system, attractive is that it appears to tie in with evolutionary theory. A system that helps us make sense of the world epistemically would help us determine what was True and what was false. Knowing that there is a tiger behind a door we're thinking of going through, for example, is clearly a benefit if we're interested in staying alive. Thus, it seems immediately plausible that this system would be of use to us as rational agents. Determining what is right is important both epistemically and pragmatically.[84]

[84] Not everyone agrees that a system that always formed True beliefs would be useful. Christopher Stephens (2001) notes that Stephen Stich, for example, thinks that from an evolutionary perspective it might be better to form beliefs that help you avoid, even potential, danger, rather than form beliefs that

As Kim Sterelny (2003) notes:

> True beliefs are a "fuel for success"; they form an information store about the world that advantages the animal in many different actions, but they are not tied to specific behaviours. (Sterelny, 2003:31)

As mentioned above, the aim here is not to prove that this system *is* in fact an evolutionary construct. Cognitive psychology is enormously complicated, with a wide range of debates and perspective in their own right. As Daniel Dennett (2004) notes in his review of Sterelny's (2003) book on the evolution of human cognition:

> One cannot expect to use any particular vocabulary to describe 'the mind' (itself an embattled concept) without provoking a sizeable cohort of objectors prepared to marvel at the fact that one is still in the thrall of that deeply problematic way of describing things. (Dennett, 2004:721)

To become engaged in debates about cognitive psychology, therefore, would take me off target, ultimately detracting from the aim of my project. The claim I want to defend here is given that this doxastic system *could be* an evolutionary construct it lends weight to the claim that something like a doxastic system exists and manages our beliefs.[85]

---

are True (Stephens, 2001:161). There are two things to note here. The first is that while it may be better at times to form a false belief because of its benefit in keeping us alive I take it that this would not always be the case. It seems that determining how the world is, even cautiously, would require a fair-degree of Truth-tracking. As Helen De Cruz, Maarten Boudry, Johan De Smedt and Stefaan Blancke (2011) argue, the "evolutionary approaches to human cognition must at least allow for approximately reliable cognitive capacities. Approaches that portray human cognition as so deeply biased and deficient that no knowledge is possible are internally incoherent and self-defeating" (De Cruz, Boudry, De Smedt & Blancke, 2011:517). The second is even if evolution means that our doxastic system is poorly adapted to track the Truth the system being discussed here is still a system that governs our beliefs. Thus, while a doxastic system may prefer something like caution rather than Truth in governing the processes of belief this still postulates a system that governs our belief processes, which in terms of the argument in this section is all I need.

[85] For an overview of some of the debates around evolutionary theories of cognition see *The Evolution of Cognition* (2000) edited by Cecelia Heyes and Ludwig Huber and Kim Sterelny's (2003) *Thought in a Hostile World: The Evolution of Human Cognition*. David Papineau (2004) describes Sterelny's book as "the first book I would recommend to any graduate student who wants to learn about comparative and evolutionary approaches to cognition. Sterelny provides an excellent introduction to nearly every topic of current philosophical interest in this area" (Papineau, 2004:491). For another discussion on Sterelny's book see Andy Clark's (2005) review.

My idea that a process of evolution may have evolved a doxastic system follows Chomsky's (1959) theory of language acquisition. Chomsky argued that within the mind/brain are certain structures that provide the foundation for us to understand and develop language (Swartz *et al*., 2004:230). As Leslie Swartz, Cheryl De La Rey and Norman Duncan (2004) explain:

> Noam Chomsky (1959) ... revolutionised psycholinguistics with an alternative view of language acquisition. Arguing that the capacity for language is not learned but rather innate ... Chomsky hypothesised the existence of a *language acquisition device*, and described humans as genetically pre-programmed to learn language. (Swartz *et al.*, 2004:230)

Similarly, the suggestion I'm making is that something like the "language acquisition device" (*ibid*.) exists with reference to beliefs.[86] The idea here is that we have an innate structure in our minds that provides the foundation for us to form, revise and retain those propositional attitudes that we think are True (beliefs).

There are two reasons why I think we might think that something like the doxastic system is plausible as an evolutionary construct or pre-existing structure in our minds/ brain. The first is that many of our beliefs are formed and revised unconsciously. As I noted in Chapter 2, many of our beliefs function outside the scope of full consciousness. In addition to having beliefs that operate in the background, or are not being viewed by the agent in full consciousness, many of our beliefs are formed and subsequently revised unconsciously. This capacity to unconsciously sort epistemic

[86] This psycholinguistic viewpoint was developed by Noam Chomsky (1959) in response to several criticisms of the existing theories of language acquisition (Swartz *et al.* 2004:230). In particular, Chomsky didn't feel that other theories were able to make sense of several common features of how children used and acquired language skills. Now, my hypothesis, similarly, suggests a doxastic system in virtue of the fact that the existing theories of the belief-truth relationship fail to account for several descriptive features of our beliefs. However, while Chomsky's theory was aimed against other theories of language acquisition my theory isn't aimed at theories of belief acquisition. To be honest, I have no particular preference for any of the theories of belief acquisition. I'm happy to acknowledge that we acquire beliefs, without, at this point, being interested in how we do so. I'm using Chomsky as a heuristic to find an alternative way of understanding the belief-truth relationship. I will not be defending this doxastic system against other theories of beliefs acquisition. For a good starting point for theories of belief acquisition, The Stanford Encyclopedia of Philosophy entry on "Belief" sketches out some of the moves in this debate.

from non-epistemic reasons, revise for truth, and so on, suggests that something like the doxastic system is in place helping us determine what is True or Truth-bearing and what is false, in much the same way that Chomsky's (1959) language construct helps people acquire language skills.[87]

Secondly, the notion that there are pre-existing structures that govern our mental lives is not a new one. Chomsky's (1959) language device is just such a structure. If we accept Chomsky's suggestion that there is a pre-existing structure in our minds that helps us with language it is no great leap to think that we have a similar construct in our minds/brains helping us make epistemic sense of the world. In addition to Chomsky's language structure, others such as Sterelny (2003) have argued that humans have developed evolutionary "capacities to track features of the environment where that tracking capacity does not drive a specific behavioural response" (Sterelny, 2003:4).

These capacities to determine how the world *is* look similar to the type of doxastic system that I'm proposing. What's important here is that these capacities that Sterelny (2003) is talking about aren't interested in determining what the world is like in order to motivate some action, but instead in order to determine how the world is, *tout court*. This suggests a structure that is interested in determining what to take as True for epistemic rather than instrumental reasons, which ties in rather neatly with my suggestion of a doxastic system.

[87] It could be argued that the doxastic system is not a pre-existing structure in our minds. That instead of having a doxastic system what we have instead is something like doxastic muscle memory: forming, revising and retaining beliefs as a result of long habituation. If this unconscious ordering of beliefs is as a result of a habitual process, then we don't have an innate doxastic system. I suspect that whether the doxastic system should be considered pre-existing or as a developed system through a continuous process of organising beliefs hinges on whether one thinks that babies have, or form beliefs, retain beliefs and so on. If they do, this suggests that something like the doxastic system, like Chomsky's (1959) language acquisition system, is a pre-existing structure in our brains/ minds. If they don't, if beliefs are mental states we can only be said to have once we've understood the concept of belief, then it nevertheless remains possible that over a long period of time in which we start to form beliefs, looking at evidence for and against, that we start to do things automatically. Of course, if this unconscious following of doxastic norms through habituation is the correct way to understand beliefs, then the doxastic system is no longer an innate construct in the way that Chomsky's language acquisition construct is innate. My gut feel is that agents start to engage in doxastic behaviour (forming beliefs and following doxastic norms) long before they can fully appreciate notions like epistemic reasons being the only reasons we can take as a reason for belief and so on.

In addition, Sterelny (2003) suggests that while there is a great deal of controversy about "the internal architecture of the human mind … at minimum … the idea [is] that the cognitive system of human agents has two subsystems. One has the function of registering states of the world as it is and as it might be. Another functions to register and rank the way the world could be changed" (Sterelny 2003:6). Again, this former subsystem seems to contain as part of it the capacity to make sense of the world epistemically.

Jesse Bering (2011), meanwhile, argues that our minds/ brains are pre-programmed to have certain beliefs. Bering, an atheist psychologist, argues that "we are programmed to believe in God" (Bering, 2011). As he puts it:

> Many scientists believe human beings evolved a suite of cognitive traits that are more or less unique to our species … [And it] is instinctive for us to seek a grand, moralistic mind that is not there. God is the default stance. (Bering, 2011)

Bering's (2011) argument suggests not only the existence of a doxastic system, but in fact a system that causes us to have certain, specific beliefs. Now, my intention is not to defend Bering, but rather to highlight that the notion of a doxastic system is not so farfetched as to be unimaginable.[88]

It seems clear that something like a doxastic system could provide an evolutionary advantage. In addition, while my theory may be scientifically flawed I take comfort in the fact that others such as Sterelny (2003) and Bering (2011) appear to endorse something similar.

As I've been at pains to point out my objective in this section was not to prove that the doxastic system in fact exists. The aim was to show that the postulation of a doxastic system is plausible. Should we find that something like a doxastic system is compatible with evolutionary theory it provides us with a reason to think that something like a doxastic system, rather than an inherent link between beliefs and

[88] For a full account of Bering's (2011) argument see Bering's *God Instinct: The Psychology of Souls, Destiny and the Meaning of Life*.

Truth, provides the explanatory power needed to explain why beliefs seem compelled to follow epistemic norms.

I believe that if we accept Chomsky's (1959) idea of a system that helps us acquire language and can see the merit in the suggestions made by Sterelny (2003) and Bering (2011) I have achieved this objective.

In the following section I present an argument for why we might think that we have a doxastic system that's independent to evolutionary concerns.

3.3. Doxastic systems and their explanatory power

Another reason to think that something like the doxastic system exists is that it allows us to keep a Tensionless Account of the truth-belief relationship. If a doxastic system can provide the explanatory tension needed, then what makes postulating beliefs as part of a doxastic system attractive is that allows us to keep beliefs as things that we think are True, *simpliciter*. Because my model sees beliefs as part of a wider doxastic system that bears the brunt of the explanatory work, we are able to keep the relationship between beliefs and Truth in this tensionless relationship, which explains a number of features of beliefs in an uncomplicated manner.

While this may seem like a weak reason, I suggest that it's no weaker than claiming that beliefs have a teleological relationship with Truth or a normative relationship with Truth. Both of the other accounts see their theory of the truth-belief relationship as the right way to understand the truth-belief relationship in virtue of the explanatory power of the model in question. That a doxastic system allows us to hold onto a Tensionless Account, with its explanatory power, provides us with as much reason to think that beliefs are part of a doxastic system as explaining transparency gives us a reason for thinking that beliefs are governed by a norm of correctness.

In addition, I argue below that a doxastic system could account for the features of belief that a Tensionless Account cannot, in and of itself, make sense of. As a result, we have two reasons to think that a doxastic system is part of the explanation of the truth-belief relationship. The first is that it allows us to hold onto a Tensionless

Account, which, along with being able to explain several features of belief, saves what it is to be a belief. The second is that this doxastic system can make sense of the other features of belief that are left out by a Tensionless Account. The combination of these two parts of my account of the truth-belief relationship, therefore, do a better job of explaining the features of beliefs than the other accounts we've examined.

In addition to the explanatory power I promise my account offers, I think there's another reason to think that beliefs are part of a doxastic system. It seems to me that the relationship my belief has with Truth is just one part of what it means to form, revise or retain a belief. When trying to determine how the world is my perceptions, memories, intuitions, gut-feelings, other beliefs about how the world was, is, and may be all seem to play a part in how I determine how the world is now. Thus, in determining what is True – in forming beliefs, revising beliefs and retaining beliefs – a wide range of mental states appear to come into play.

While other theories of the truth-belief relationship don't necessarily preclude this, by suggesting that what governs our belief processes is the relationship between the individual belief itself and Truth it does seem to undermine the impact these other considerations might play in our doxastic lives.

What makes the postulation of a doxastic system attractive from my perspective is that it acknowledges that explaining all the features of belief is complex. That perhaps it is not solely the relationship that a belief has with Truth that is involved in the processes behind belief formation, revision and retention.

A third argument in favour of this system is that in addition to other cognitions having an effect on our beliefs, beliefs have an effect on these other mental states. This suggests that beliefs and other mental states are closely linked.

I've suggested above that a doxastic system would include not only our beliefs, but in addition, all other cognitions, conations and feelings that pertain to tracking the Truth. These would include, but are not limited to, our perceptual faculties, our ability to reason, our reasoning skills, our intuitions, hunches, desire to determine the Truth, and the ability to appreciate evidence and counter evidence. If I'm right, and beliefs are

part of a doxastic system, then in addition to perceptions, reasoning skills and so on impacting on our beliefs (which would seem part of the natural process of evidence gathering to form a belief), we should expect to see what we believe impact on, for example, what we see.

While this may offend our epistemic sensibilities, this is precisely what appears to happen. Larry Samovar, Richard Porter and Edwin McDaniel (2010) note that our beliefs shape the way we view the world, not only in terms of values, but in terms of what we see (Samovar, Porter & McDaniel, 2010:184-185).

> The moon is a rocky, arid physical sphere that orbits the earth; yet, when looking at this objection, many Americans often visualize a human face; many American Indians, as well as Japanese, perceive a rabbit; Chinese claim to see a lady fleeing her husband; and Samoans report a woman weaving … The reason [different cultures see the same object differently is that] … diverse cultures have taught their members to look at the world in different ways. (*ibid*.)

What this appears to highlight is that beliefs don't operate independently of our other cognitive states, rather they appear to inform and be informed by these states. While we don't need a doxastic system to see that beliefs interact with other mental states, that perception can affect beliefs and *vice versa* suggests that beliefs as well as other cognitions interact with one another quite intimately.

The other models of the truth-belief relationship we've looked at postulate beliefs themselves as having some sort of relationship with Truth. This relationship between just this type of mental state and Truth does two things. The first is that it minimises the impact that other cognitions, feelings and conations have on our beliefs. This wouldn't be a problem if our beliefs managed to always take what was epistemically relevant from our other mental states in forming our beliefs. However, this doesn't seem to be the way it works. The way we see the world appears to be impacted by more than just truth-relevant considerations. My dislike for an interlocutor may make me disregard the epistemic value of his/her remarks. The second, following on from this first reason, is that sometimes the Truth of our beliefs is threatened or completely

undermined by the non-epistemic reasons in our other mental states. It is this that causes us to form motivated, biased, incorrect and evidence-insensitive beliefs. What this suggests is that beliefs don't just have a relationship with Truth and the Truth-relevant bits of other mental states, they have a relationship with our mental states.

While this doesn't prove that there is such a thing as a doxastic system, accounts of the truth-belief relationship that fail to take into account our other mental states and the impact they have on one another struggle to make sense of when beliefs seem to fall away from Truth. Given that we do see beliefs interact with other mental states – sometimes mistaking these mental states as relevant epistemic reasons for a belief – suggests that beliefs are part of a wider system and aren't quite as separate as other models of the truth-belief relationship suggest.

Thus, there are three, non-evolutionary, reasons to think that we have a doxastic system.

One, it allows us to keep a Tensionless Account with its concomitant benefits. Further, if I can show that a doxastic system is able to provide the required explanatory tension, we now have a model that is able to provide an account of all of the features of belief I discussed in Chapter 2.

Two, a doxastic system allows us to make sense of the influence of our other mental states on our beliefs and *vice versa*.

This in turn provides us with the third reason to think that beliefs are part of a doxastic system. A model that explains how other mental states impact on our beliefs provides an uncomplicated explanation of motivated and biased beliefs. What happens in cases of motivated beliefs is that the non-epistemic reasons in question (interacting with our beliefs) cloud our judgement, which causes us to mistake non-epistemic reasons as epistemic ones.

While it may seem unnecessarily complicated to suggest both that beliefs have a tensionless relationship with Truth and are part of a system that controls our beliefs, the explanatory power that postulating this move has stands in its favour.

These reasons for thinking that something like a doxastic system exists, however, don't explain how this system would provide the explanatory tension required to explain why beliefs appear to be compelled to follow epistemic norms. In the section below I propose an argument for how the doxastic system is able to provide the required explanatory tension.

4. The Problem of Motivation

One of the distinctive features of beliefs is the fact that they regulate for truth. In doing so, beliefs appear to follow certain epistemic norms. Among them: only being able to take on board epistemic reasons as reasons for belief; and then regulating in the face of these reasons, often unconsciously, should we find these reasons epistemically sufficient. Thus, the doxastic norms are so strong as to ensure that an agent seemingly always follows epistemic reasons in forming a belief.

David Owens (2003) notes that there are two things that an account of the truth-belief relationship needs to account for "the authority of epistemic norms [and] … also their motivational efficacy" (Owens, 2003:284).

The first component is that epistemic reasons are the only reasons an agent can see him/herself as having for a belief. And furthermore, epistemic reasons are the only reasons an agent can take on board as reasons for or against a belief. This, as has been noted in Chapter 2, is significant. No matter how badly an agent wants to take some non-epistemic reason as an epistemic reason he/she cannot.[89]

[89] This is not to say that I don't form beliefs based on non-epistemic reasons. Instances of motivated beliefs are precisely a case in which I form a belief for non-epistemic reasons. What is being denied is that I can see a non-epistemic reason as a reason to believe something. Thus, I can't knowingly take my admiration for Napoli manager, Rafa Benitez, as a reason to believe Napoli will win the Serie A football title this season. It is also worth noting that things that are traditionally thought of as non-epistemic reasons (such as emotions) can be epistemic reasons. In the simplest of cases, feeling anger towards someone counts as an epistemic reason for believing I am angry with someone. What is being denied is that a non-epistemic reason can count as a reason to think something is True. If my feelings of anger towards someone somehow make me believe that what he/she is saying is false, I have formed a belief based on non-epistemic reasons. The claim above is not that we can't do this, but that we can't think of one of our current beliefs that the reasons I believe that *p* is based on or sustained by some non-epistemic reason.

The second component is that we are compelled to act in accordance with these epistemic reasons.[90] Thus, not only can I not choose to take some non-epistemic reason as an epistemic one, I also cannot ignore epistemic reasons even if I really want to.[91]

Being able to explain these two components (what I call the Problem of Motivation), is essential for any account of the truth-belief relationship.[92] So, how, under my account, is it that we are compelled to follow these epistemic norms?

In what follows I show how a doxastic system that has the goal of Truth could provide the explanatory tension to explain the Problem of Motivation.

Having done so I argue that while I think a doxastic system must, in virtue of being a doxastic system, have a goal of Truth this goal is tempered by the need to navigate and take on board other cognitions, emotions, conations and dispositions. The doxastic system can be said, therefore, to *de facto* aim at rationality.

4.1. Doxastic systems and explanatory tension

In this section I argue that a doxastic system with the goal of Truth can provide the explanatory tension needed to explain the Problem of Motivation. In doing so, this

[90] It is important to note that not every epistemic reason will have this impact on our beliefs, but only what I take to be sufficient epistemic reasons. Being told that my partner is having an affair is an epistemic reason to believe that my partner is having an affair. However, in this case it may not be sufficient to cause me to suspend my belief that my partner is faithful or form a belief that she is unfaithful. I may not, for example, trust the person who told me. What counts as sufficient epistemic reasons to form or suspend a belief will be different from person to person and possibly from belief to belief.

[91] I might strongly want to ignore all reasons that point to a corrupt government, for example, but faced with continuous evidence for the Truth of this claim I can't help but believe it. Similarly, I cannot, again *ceteris paribus*, ignore evidence against a belief I do have, no matter how badly I want to hold onto the original belief. The cuckolded husband, in the face of evidence of his partner's infidelity, believes that he has been betrayed no matter how much he'd prefer not to have that belief. Epistemic reasons motivate us to "act" (to believe that *p* or to stop believing that *p*) in a way that is so powerful that it is traditionally extremely difficult to fail to "act" in accordance with those reasons, irrespective of one's desires.

[92] Steglich-Petersen (2006), for example, argues that Normative Accounts fail because they cannot get round the Problem of Motivation (Steglich-Petersen 2006:506-507). If Steglich-Petersen is correct, and a Normative Account can't make sense of the Problem of Motivation, then these accounts should be abandoned in favour of something that can make sense of the Problem of Motivation.

system is able to account for most of the features of belief that a Tensionless Account leaves out, barring *idées fixe*.

One of the reasons that teleological accounts are so attractive, whether they are Aim or Purpose Accounts, is the fact that they provide a solution to the Problem of Motivation.

> Postulating a goal for belief explains not only the authority of epistemic norms but also their motivational efficacy. And, in the eyes of many philosophers, this is essential for epistemic norms to have any authority. (Owens, 2003:284)[93]

What's not in the quote, but needs to be mentioned, is that the goal here must be Truth or something strongly linked with Truth.[94] If, for example, beliefs aim at pragmatic usefulness, then we could not explain the Problem of Motivation. If beliefs aim at pragmatic usefulness, then there will be times when the means we follow to achieve this end are not epistemic. For if believing that *p* would be instrumentally useful but not True, then a belief aiming at pragmatic usefulness would not take on board only epistemic reasons and would not be motivated to act on them. Thus, if beliefs or a doxastic system aimed at instrumental usefulness it could not get round the Problem of Motivation.

If, however, beliefs have a constitutive goal of Truth, then the only way to achieve that goal and have a belief is to be necessarily motivated by epistemic reasons and only epistemic reasons. Failure to take on board epistemic reasons and act in

[93] I've suggested in Chapter 3 that Owens (2003) cashes out the teleological claim as essentially an Aim Account, rather than as a Purpose Account. However, as I argued in Chapter 3, a Purpose Account is equally able to provide the necessary explanatory tension in order to resolve the Problem of Motivation.

[94] As I've mentioned in my introduction some of the recent literature in this debate has shifted the debate from the truth-belief relationship to a relationship that beliefs have with knowledge. Given that knowledge appears to be sufficiently close to Truth in terms of the demands this aim, purpose or norm would generate it could be supposed that knowledge, as an aim, purpose or norm, would be able to solve the Problem of Motivation in similar ways to Truth. For the sake of clarity and time I've avoided engaging with this literature, focussing solely on how the truth-belief relationship could/should be cashed out. My objection to the truth-belief relationship has been in terms of thinking that beliefs themselves aim at, have the purpose of or themselves directly follow the norms of Truth and my suggestion is that swapping Truth for knowledge won't necessarily solve these problems. For an alternative objection to knowledge accounts see Daniel Whiting (2013).

accordance with them would be to fail to adhere to the goal in question. Thus the primacy and motivational force that epistemic reasons have over our beliefs is explained in virtue of the constitutive goal of belief.

As I've argued, more or less *ad nauseum*, the problem with existing teleological accounts is putting Truth and belief in the wrong relationship with one another such that what it means to be a belief is lost. To put beliefs into the right relationship with Truth I've suggested removing the explanatory tension. In doing so, I've saved what it's like to be a belief, but created a problem in terms of explaining the Problem of Motivation.

Now, my suggestion is that the doxastic system, of which beliefs are just one part, has the goal of Truth. All the things that make up this doxastic system – things like reasoning skills, perception, other beliefs and so on – would operate together in order to discharge the system's constitutive goal: Truth.

Could this system account for the Problem of Motivation? The answer seems to be an uncomplicated, yes. In much the same way that both Aim Accounts and Purpose Account explain why we can only think that epistemic reasons are reasons for a belief, and in much the same way that these accounts highlight why these reasons would directly impact on that belief, causing us to act in line with them, a doxastic system with a goal of Truth would demand certain things that would be generated as a result of the goal of the system.

Thus, if the constitutive goal of a doxastic system is Truth, then the only way to meet this goal would be to ensure that epistemic reasons and only epistemic reasons are taken on board with regards to beliefs. Further, it would require that these reasons motivate the system to act on them. In the case of a doxastic system, this would involve all of the constitutive parts of the system working together to ensure that the system achieves its goal. Thus our vision and hearing, for example, might work together to determine how far away an approaching car is and when it might reach us. In doing so, as a consequence of this system's goal of Truth, beliefs would be formed.

Further, a system with a goal of Truth would need to ensure that all of its constitutive components were operating in such a way as to ensure that they were maximising the chances of the system achieving its goal. Thus, if we're wary that what we are seeing might not be True (we're watching an illusionist, for example), other components in the doxastic system would be working to ensure that we don't form a belief solely on the perceptual evidence. In this case, our background beliefs about illusionists, for example, would inform our perceptions and *vice versa*. This system, with the components working together, would help us separate the true from the false, helping us determine how the world actually is and, as a result, form veridical beliefs.

One of the problems with both Aim Accounts and Purpose Accounts was in order for these accounts to provide any explanatory tension they had to create a gap between beliefs and Truth and that this lost what it was like to be a belief. However, in the case where it is the system itself that has the goal of Truth, there is no need to have an explanatory gap between beliefs and Truth. Or, more accurately, the gap between the two is part of two different things rather than being shoehorned into the same mental state. This allows us to explain the tension in virtue of the goal of the system, while beliefs are formed as part of this system's attempt to meet the goal of Truth. Thus, we can think that something is True (in a way that cashes out what it is to believe) and still have the explanatory tension that would account for why beliefs regulate for truth. The beliefs themselves do not provide the force that regulates them, instead they are regulated as a result of the system pursuing its constitutive goal.

To see how this works, consider how the homeostatic system regulates a body's temperature. The homeostatic system's job is to regulate our body temperature. The processes that are involved in this system are such that should our temperature exceed a certain threshold certain processes are activated to try to bring the temperature of the body down. Similarly, should the body temperature drop below a certain threshold another set of processes are activated to try to bring that temperature back up to an acceptable level. Thus, we sweat when we get too hot and shiver when we get too cold. In the case of the homeostatic system, this system has the goal of regulating our body temperature and, as a result of that purpose, the various processes could be said to "aim" at fulfilling this purpose. However, in this case, unlike the case of an archer aiming at a target, the so-called aims are part of the inherent goal of the homeostatic

system. What does the work in the homeostatic system is the inherent purpose of that system.

Similarly, I'm suggesting that what drives the motivational force of epistemic reasons is the inherent purpose of the doxastic system as a whole. A belief, therefore, could be said to aim at the Truth, but they only do so in the same way that sweat aims at cooling down our body as part of the doxastic system's inherent goal. Beliefs, like sweat or shivering, are the result of the system operating to discharge its function.

Now, consider a belief being revised under this doxastic system. The system and its processes aim at trying to fulfil the inherent process of the system, in much the same way that the homeostatic system and its processes aim to fulfil that system's goal. Should the body fail to be at the correct temperature the system reacts to try to regulate this system. Similarly, should the doxastic system encounter counter evidence, highlighting that the inherent goal of the system hasn't been met, the system kicks in to regulate the system. In this case, jettisoning the egregious beliefs, adjusting the way we perceive the world, possibly regulating our emotional responses, drives, and so on. A doxastic system, with the goal of Truth, therefore, seems well placed to provide the explanatory tension required to explain how beliefs follow doxastic norms.

Combined with a Tensionless Account we can not only explain why only epistemic reasons count as a reason for belief we can also explain their authority and motivational force. As a result, we can explain how it is that beliefs regulate for truth. And between the two parts of this account of the truth-belief relationship we are, therefore, able to make sense of the majority of the features of belief.

In terms of the Tensionless Account we can make sense the direct link between beliefs and truth as well as the scope of belief. In addition, we have an explanation for why, if not how: only epistemic reasons count as reasons to believe; we have false beliefs; we have biased beliefs; we have beliefs based on inconclusive evidence; we have *idée fixe*; and beliefs regulate for truth.

In terms of answering the *how* question, a doxastic system with the goal of Truth is able to answer the Problem of Motivation, which, in turn, explains how beliefs regulate for truth.

Above I argued that a doxastic system wouldn't operate in isolation. As a result, we could expect at times for the system to mistake inconclusive or non-epistemic reasons as legitimate reasons for belief. In much the same way that anxiety can trick the body to react as if it is cold (causing us to shiver), the suggestion here is that sometimes other parts of our mental functioning can "trick" the doxastic system into forming a belief. Thus, we have an account of how motivated beliefs; false beliefs and instances of biased beliefs can occur. It must be noted, however, that a doxastic system with the sole goal of Truth, with the entire system aiming solely and exclusively at Truth, cannot account for *idées fixe* or for beliefs based on inconclusive evidence.[95]

That said, my model nevertheless has a number of advantages over the existing models we've looked at.

By postulating a tensionless relationship with Truth it leaves belief in the "right" relationship with Truth, which in turn explains a number of features of belief. In virtue of suggesting a constitutive goal for the system as a whole, rather than with belief itself, this model solves the Problem of Motivation – explaining both the authority and motivational force of epistemic norms.

[95] I argued in Chapter 3 that if Purpose Accounts were correct then we should see a belief being constantly pulled towards correctness. What I didn't do was examine Purpose Accounts' ability to make sense of *idées fixe*. It seems clear to me that Purpose Accounts cannot make sense of *idées fixe*. Given a Purpose Account, what we should expect to see under a Purpose Account is *idées fixe*, at the very least, being suspended in the face of sometimes overwhelming counter evidence. If beliefs are being pulled into correctness, then being faced with overwhelming evidence for the truth of a claim should see our beliefs at the very least suspend in the face of that evidence. Since some of our beliefs do maintain in the face of counter evidence Purpose Accounts aren't the right way to understand *idées fixe* and are subsequently undermined as the correct way to understand the truth-belief relationship. Similarly, if a doxastic system has the purpose of Truth in the same way that a Purpose Account sees beliefs has having the goal of Truth, then it seems that this doxastic system could not account for *idées fixe*. In this case the system as a whole would work together to pull the belief towards correctness. While we might be prepared under this model to grant a little more time and leeway to beliefs – allowing for the system as a whole to regulate for Truth – in terms of how long they remain evidence resistant we should still expect to see these beliefs, at the very least, suspend in the face of this counter evidence. That they don't (or don't always) means that a doxastic system with the sole goal of Truth cannot explain all the features of belief.

However, despite the attraction of having a doxastic system with the purpose of Truth, I don't think that this is the way it works. For one thing, as it stands, this account cannot make sense of *idées fixe* and I take this as a serious problem.

In the following section I argue that if a doxastic system exists it must have the primary goal of Truth. Given the benefits of a model of the truth-belief relationship that has a Tensionless Account operating as part of a doxastic system with the goal of Truth I examine an argument that could, using this model, explain *idées fixe*. I argue, however, that this argument fails to explain *idées fixe* to my satisfaction.

I then suggest an alternative solution, which I believe ties in with how we believe. I argue the primary goal of Truth is tempered by other factors that the doxastic system has to, or does, incorporate when trying to reach its goal. Taking these other considerations into account means, I argue, that the doxastic system *de facto* aims at rationality.

5. Beliefs and aiming at rationality

I argued above that a doxastic system with the goal of Truth would account for the Problem of Motivation. What I didn't do, however, was argue why a doxastic system would aim at Truth rather than say psychological wellbeing. The reason is that if such a system exists, it must, in virtue of what it claims to be (namely a system that governs our beliefs), have the goal of Truth.

To be a belief is to take it that some proposition is True. Not for any extrinsic reason transparent to you, but because you take the proposition as actually being True. Thus, if I have a system that is trying to form True cognitions this system must have Truth as its goal. To have some other goal would be to aim at forming, revising or retaining some other mental state. Now, I'm not denying that other structures do or might exist. In fact I've suggested that what I'm calling the doxastic system would interact with other systems or subsystems as part of an agent's mental functioning as a whole. All I'm noting is that if a doxastic system exists, and I've argued we have good reason to think it does, the primary goal of this system, because it deals with beliefs, must be Truth.

If we can find a way to explain *idées fixe* and beliefs based on inconclusive evidence we have a model of the truth-account relationship that makes sense of all of the features of belief.

Ryan McKay and Dennett (2009), who take an evolutionary perspective on how we form false or incorrect beliefs, may have a solution for how to explain *idées fixe*, while keeping a doxastic system with the goal of Truth. They ask:

> From an evolutionary standpoint, a default presumption is that true beliefs are adaptive and misbeliefs [motivated or false beliefs] maladaptive. But if humans are biologically engineered to appraise the world accurately and to form true beliefs, how are we to explain the routine exceptions to this rule? How can we account for mistaken beliefs, bizarre delusions, and instances of self-deception? (McKay & Dennett, 2009:493)

Their answer to how it is that we form motivated or false beliefs is that firstly, we oughtn't to think "that evolution is a perfect designer" and secondly, there may be cases "where the truth hurts so systematically that we are actually better off with falsehood" (McKay & Dennett, 2009:509).

> We have seen that in such circumstances falsehood can be sustained by evolved systems of misbelief. So, in certain rarefied contexts, misbelief itself can actually be adaptive. (*ibid*.)

Now, while McKay and Dennett (2009) are trying to explain incorrect beliefs generally, I take it that we can use the same strategy to explain why some of our beliefs maintain in the face of counter evidence. In some cases these beliefs are the result of a malfunctioning doxastic system, while in others, "where the truth hurts so systematically that we are actually better off with falsehood" (*ibid*.), our system stops us from regulating the belief. In both cases we could still say that the system had the goal of Truth, just that given the evolutionary story behind the system sometimes it fails and sometimes other considerations are more important.

I've suggested that there are four different types of *idée fixe* we need to make sense of:

1. Bigoted beliefs;
2. Delusions;
3. Things I am going to believe no matter what the argument against them are;

And,

4. Beliefs that maintain in the face of unanswerable, yet unbelievable arguments.

Using McKay and Dennett's (2009) suggestion, we are able to say of delusions, for instance, that they are cases in which the system has broken down and that in the case of beliefs in the face of unanswerable and yet unbelievable arguments, for example, these beliefs provide us with an evolutionary advantage. While this solution is appealing I still think that my claim that a doxastic system aims at rationality is a better fit for the doxastic phenomena we've examined in this project.

I briefly sketch out my worries with McKay and Dennett's (2009) suggestion before I turn to my own positive account. The worries are by no means knockdown arguments, but concerns that I think my account is able to avoid. My argument, briefly, is that there are some potential worries to McKay and Dennett's solution that my account is able to avoid. In addition, I believe that my account provides a non-evolutionary perspective that makes sense of what it is like to believe. Therefore, it provides a less-complicated or more-appealing way to understand the truth-belief relationship as a whole.

In terms of my worries with the proposed solutions above: if we take McKay and Dennett's (2009) suggestion on board it requires us to buy into an evolutionary story about beliefs and mind. There are several potential pitfalls here. The first is that this evolutionary theory could be wrong. The second is having distanced myself from nailing my argument for a doxastic system to an evolutionary flag this would put me firmly back in need of explaining this system from an evolutionary perspective. The third is that it looks like we're going to end up with a completely naturalistic, potentially reductionist picture of the mind, which I, for one, am not comfortable with.

In addition to these evolutionary concerns, it seems to me that some beliefs that maintain in the face of unanswerable and yet unbelievable arguments are neither the result of a breakdown of the system, nor provide us with an evolutionary advantage.

I have an atheist friend, for example, who finds the ontological argument for God's existence both unanswerable and yet unbelievable. If Bering (2011) is right, then it is a belief in God that gives us an evolutionary advantage. Assuming the truth of both of these claims, how are we to understand my friend's belief? Given that it's not a belief that provides an evolutionary advantage it must be, according to McKay and Dennett (2009), as a result of a breakdown in the system. But this doesn't seem right. While we might be happy to talk about delusions as a breakdown in the doxastic system (one of the reasons I find McKay and Dennett's suggestion appealing), it seems odd to say of my friend's belief, which is based on a number of (possibly unsatisfying, given the ontological argument, but nevertheless) epistemic reasons, that it is as a result of a breakdown of the system.

While this is the specific case, the more general worry I have is that some *idées fixe* appear to neither a) represent some sort of doxastic breakdown, nor b) provide us with any kind of evolutionary advantage. It seems in these cases we would need another explanation for the belief and its evidence-resistant nature.

My account, I believe, can make sense of these beliefs, without needing to appeal to the move made above. In what follows I suggest that a doxastic system that aims at rationality is able to make sense of my personal doxastic bugbear – evidence-resistant beliefs – in a way that I feel makes more sense than McKay and Dennett's (2009) suggestion.

5.1. Rationality

To make a rational decision is to, roughly, make a decision that "makes the most sense" all things considered given the specific end. As Richard Foley (1987) puts it, rationality *qua* rationality is pursuing the most effective means to an end, "all else

being equal" (Foley, 1987:6). Thus, if your goal is Y and the most effective means to Y is X, it is rational to do X.

Because being rational is dependent on the goal in question, being rational will differ from situation to situation dependent on the agent's goals. The subject in Barbara Winters' (1979) experiment where the agent "is truthfully informed that either a gentleman or a hungry tiger waits in the next room and that the outcome has been decided by the concealed flip of a fair coin" (Winters, 1979:250) would be making a rational decision, assuming she's interested in staying alive, if she went into the room armed with a loaded revolver. This is not to say she thinks it True or even that it is True that there is a tiger in the room. She may, in fact, be fairly certain that there isn't a tiger in the room. However, given that the consequences of getting it wrong would impact quite dramatically on her goal of staying alive it would nevertheless be rational to take the loaded revolver into the room.

If, however, Winters' (1979) character is interested in being mauled by a tiger (she has decided that this is the way to end her life or at least get a fairly story-worthy set of scars) then it would be rational to a) hope for a tiger behind the door and b) not take the revolver with her in case she accidentally shoots the tiger.

As with instrumental rationality above, epistemic rationality, what it is rational to believe, depends on what we define as the epistemic goal. Foley (1987) thinks the epistemic goal of doxastic agents is to believe things that are True and to not believe things that are false (Foley, 1987:8). However, as I've been at pains to point out in Chapters 2 and 4 the goal of a doxastic agent, with regards to belief, needn't be Truth, and is incidental to the question of what the goal of a doxastic system might be.[96] What I'm interested in is not the goal of the doxastic agent, but rather the goal of the doxastic system.

[96] In Pascal's Wager the agent sets him/herself the goal of believing that God exists, irrespective of the Truth of the claim. Such a person, pursuing this goal, could be said to be rational if he/she found effective means to reach this goal. In this case, deliberately misleading him/herself could and/or would be rational. The same, of course, could not be said of someone who was trying to determine the Truth of God's existence. In this case, the only rational actions would be Truth-directed ones.

I've argued above that the primary goal of the doxastic system must be Truth, but that if this system pursued an untainted goal of Truth, then we can't account for *idées fixe*. Such a system, behaving completely rationally, pursuing a goal of Truth, *simpliciter*, all things being equal, cannot make sense of the fact that some of our beliefs maintain in the face of sometimes overwhelming counter evidence.

But *idées fixe* do exist. So how do we make sense of these beliefs? As I've noted above, the move that McKay and Dennett (2009) appeal to doesn't work for me. My suggestion, one that I think makes sense of *idées fixe* within a system whose primary goal is Truth, is to claim that the goal of Truth is shaped by, among others, what an agent has previously taken as Truth as well as other, perhaps non-epistemic, conative and emotional states.

Given that this system interacts with and must take on board other commitments, the system *de facto* aims at rationality. Thus, the Truth goal is tempered in virtue of needing to work around or incorporate key and entrenched cognitions, emotions and conations that an agent has and a doxastic system aims to rationally achieve this tempered Truth goal.

In the following section I lay out an argument for the claim that a doxastic system would need to incorporate and work with and around other features of an agent's mental landscape, thereby shaping the goal of Truth and causing the system to aim at rationality.

5.2. The doxastic system and a tempered-Truth goal (the aim of rationality)

There are two reasons that I think that a doxastic system aims at rationality in the way that I've suggested.

The first, following Willard Quine and Joseph Ullian (1978), is that beliefs aren't formed, retained and revised in isolation. Instead, our beliefs inform other beliefs and potential beliefs. Thus, when a doxastic system is considering evidence for something it takes these existing beliefs into consideration. Given that beliefs are, to the agent, legitimate Truth claims, these beliefs are taken as epistemic reasons for or against the

proposition under consideration for belief. As a result, our existing beliefs, some more than others, shape or interfere with a doxastic system reaching Truth, *simpliciter*.

The second is that in addition to other beliefs shaping our doxastic system, other mental states shape the way our doxastic system goes about its goal of Truth. I briefly argued above that a doxastic system wouldn't operate in isolation, but would be one part of an agent's mental functioning as a whole. Thus, we would expect to see beliefs having an impact on how one reacts emotionally towards things and *vice versa*. In what follows I explore this in greater detail. Having made these two arguments I then look at how this might affect the doxastic system.

5.2.1. A doxastic system and a web of belief

In this section I argue that in determining what is True a doxastic system must take the agent's other beliefs into consideration. These beliefs, which are the things that an agent already takes as True, act as Truth-bearing reasons. The goal of Truth, therefore, is shaped by an agent's already existing cognitions.

It seems a fairly uncontroversial claim to suggest that our beliefs don't act independently of one another. A belief in Christianity, for example, would likely tie in with the beliefs: God exists; and God is interested in us as human beings. Each of these beliefs, though strictly independent, ties in with one another such that changes in one are likely to affect the others.[97]

In addition to these links between beliefs, it seems that forming a new belief will depend to some extent on the beliefs we already have. This is what I mean when I say that a doxastic system's goal of Truth is shaped by the pre-existing beliefs an agent has. Quine and Ullian (1978) in fact go further. They argue not only that beliefs are interlinked and that existing beliefs would impact how we form new beliefs, but that our assessment of the Truth of a proposition may have less to do with the

[97] If I give up my belief that God is personally interested in me, I may give up my belief in Christianity, though perhaps not my belief in God. If I give up my belief in Christianity I may give up both of the other beliefs – though of course I may give up neither if I convert to another religion, for example.

proposition's "independent merits" and more to do with "the light of the full body of our beliefs" (Quine and Ullian 1978:16).

> A very accomplished mechanic might be able to tell something about an automobile's engine by examining its parts one by one, each in complete isolation from the others, but it would surely serve his purpose better to see the engine as a whole with all the parts functioning together. So with what we believe. It is in the light of the full body of our beliefs that candidates [potential beliefs] gain acceptance or rejection; any independent merits of a candidate tend to be less decisive. (Quine & Ullian, 1978:16)

Now, even if Quine and Ullian (1978) are only partially right, it seems that our ability to determine the Truth, *simpliciter*, will be shaped by what I already believe. A belief in God, for example, may mean approaching the world and seeing the world for how it is in a completely different way to the way an atheist would.

Consider Shahir and Shirley. Shahir is a devote Muslim. Shahir believes that when he dies his body will lie until the Day of Judgement. On the day of judgement he will, if he's been a good Muslim, go to heaven, or hell if he's been a bad Muslim. Because of his beliefs, Shahir believes that eating pork is a sin. Given the consequences of sinning Shahir does his best to avoid eating pork, always making sure that the meat he eats comes from what he believes are reputable halal butchers. Shirley, however, is not a Muslim. In fact Shirley is not religious at all and, as a result, has no objections to eating pork and no beliefs about the link between eating pork and sin.

It just so happens that both Shirley and Shahir shop at the same butcher for their meat. When it is discovered that the beef mince at the butcher shop contains pork, Shahir forms a belief that the butcher was lazy and incompetent. Shirley, meanwhile, forms no belief about the butcher at all, but thinks that contamination is probably inevitable in modern abattoirs.

What explains the two different beliefs about the butcher (or in this case belief and lack of belief) is the fact that the evidence is interpreted by both Shirley and Shahir in

"the light of the full body of [their] beliefs" (Quine & Ullian, 1978:16). The background beliefs that the two agents have shape the way they see the evidence.

In addition to this claim I want to suggest that some of our beliefs are core beliefs: beliefs whose Truth is fundamental to the way we see the world. Samovar *et al.* (2010) allude to this in discussing how culture can have a deep and lasting impact on the way we view the world. The affect of culture doesn't just affect what we think is right and wrong, but what we see and how we see it (Samovar *et al.*, 2010:48-91 & 184-207). This worldview, in turn, has a massive impact on metaphysical and moral beliefs (Samovar *et al.*, 2010:97-150) as well as our beliefs about ourselves and others (Samovar *et al.*, 2010:153-182). In each of these cases what's being highlighted is the role that our core beliefs have in determining how we see the world. We can imagine that some of these core beliefs act like a foundation of a house such that if it was removed it would dramatically undermine the entire system of beliefs.

If this is right, then it seems to me that a doxastic system's goal of Truth is tempered by the framework provided by our existing beliefs and in particular these core beliefs. What I see as Truth will be viewed through the lenses of my already existing beliefs. The reason for this is that core beliefs, as with our other beliefs, function as epistemic evidence for the doxastic system: an existing belief that *p* is self-evidently, from an agent's perspective, a reason to think that *p* is True. If *p* entails *q*, then *q* must also be correct and so on.[98]

5.2.2. A doxastic system and non-cognitive emotions

Motivated beliefs are explained by some temporary but vivid fear that causes the doxastic system to be overwhelmed. As a result, some non-epistemic reason for *p* is mistaken as an epistemic reason for *p* and a belief that *p* is formed. In these cases, once it has come to the agent's attention that his/her belief is based on something

[98] With a core belief we not only have that belief as a reason in favour of or against a candidate that agrees or disagrees with the belief respectively, we have a whole host of reasons. Since a core belief informs so many of our other beliefs, each of the beliefs based on a core belief would count as a reason for or against the adoption of, or rejection of, a new belief. Thus, the goal of Truth is tempered by our existing beliefs, premised on certain core beliefs.

other than solid epistemic reasons, what we expect to see and what we tend to see is his/her motivated belief disappear.

However, part of the claim I'm making is that not all fears, hopes, desires etc. are temporary. In some instances the cognition, emotion, conation or combination of the three is deeply rooted; affecting the way an agent sees and understands his/her world, including the way he/she forms beliefs.

Consider Candice, who has been abandoned as a child and, as a result, doesn't trust people. This fear of abandonment and the resultant lack of trust are, in this case, likely to be deeply rooted, impacting on a number of mental and behavioural functionings and the way that she lives her life.[99]

I take this as fairly uncontroversial. The things we go through as children can have a deep and lasting impact on our lives. Eric Berne (1961), writing from a psychoanalytic perspective, notes that things that have happened to us as children affect the way that we respond to things in the future (Berne, 1961:19). Donald Winnicott meanwhile argues that many of the false beliefs we have are formed as infants through a misunderstanding between ourselves and our parents (Winnicott in Rodman, 2003:264-265).[100] These experiences affect the way we form relationships with others and how we see the world.

The potentially controversial part of the claim I'm making is that in much the same way that fears and desires can cause us to form motivated beliefs, these deep-seated emotions, conations and cognitions shape the way our doxastic system works.

---

[99] A psychologist friend describes a similar case. She describes mentoring someone who has deep-seated trust issues. Now, as part of her job as a mentor her role is to understand her student in order to help him/her achieve his/her stated goal. However, because of his/her deep-seated trust issues he/she refuses to tell her anything. When I asked her about this case she said: "I consider myself to be a trustworthy person. In addition, I don't have anything to gain by breaking his/her trust. But he/she doesn't believe me". Interpreting this case through my model: the claim I'm making is that the deep-seated trust issues of the pupil are impacting on his/her ability to form a True belief. The doxastic system, aiming at rationality is forced to weigh this dispositional state of mistrust against the goal of Truth. In doing so, the system aims at rationality, such that the agent believes what he/she is able to believe given his/her other mental states.

[100] Donald Winnicott argues that babies' true selves are communicated through "the spontaneous gesture, which is the True Self in action. The not-good-enough mother fails to meet the gesture and substitutes her own, which with the infant complies with the beginning of a False Self" (Winnicott in Rodman, 2003:265).

A doxastic system, therefore, with goal of Truth would be shaped by what we already believe, our core beliefs and other deep-seated conations and emotions. These mental states would inform and shape the evidence that the agent encounters. Because these states shape the person's doxastic system, the system aims at pursuing the most effective means to the goal of Truth all things considered, where the all things considered pertains to how these other mental states shape the system's goal of Truth.

5.2.3. A doxastic system and the goal of tempered Truth

The impact of these shaping-influences means that we could expect different people, all behaving rationally, to form different beliefs even with the same evidence. Again, this must be relatively uncontroversial. In the case of Shirley and Shahir above, the same evidence is perceived and then believed completely differently as a result of the background beliefs the two of them have.

Foley (1987) makes a similar point. In discussing two agents with the same epistemic goal of Truth, but different background beliefs, he notes that they will approach the same goal differently. Each approach will be rational, given the tempered end, but nevertheless have very different end results.

> If on careful reflection a person *S* would believe that he is infallible (and hence that all of his beliefs are true) and if in addition he would believe that he is omniscient (and hence that every truth is believed by him …), then on careful reflection he would believe that the most effective means to his epistemic goal is to believe precisely what he now believes
>
> At the other extreme is the radical sceptic, who on careful reflection would believe that no means is better than any other means as a way of achieving his epistemic goal. On careful reflection he would think that the means by which most of us come to believe what we do – perception, memory, and the like – are no more likely to generate true beliefs and to discourage false beliefs than would deciding what to believe … on the basis of a flip of a coin … Since there is nothing

> that such a person on careful reflection would take to be an effective mans to his epistemic goal, there is nothing that is epistemically rational for such a person. (Foley, 1987:9)

In much the same way that the sceptic's and God-complex agent's other beliefs impact on how they would rationally pursue the same goal, the suggestion goes that the system would need to balance or weigh up Truth against these other factors in pursuing its own goal of Truth. The goal of Truth, therefore, is tempered and individuated.

One of the consequences of my position is that we will struggle to believe things that don't fit within our existing schema. Foley (1987) defends a similar idea, though he cashes his position out in terms of an epistemic agent, rather than in terms of a doxastic system.

> A proposition *p* is epistemically rational for an individual *S* just if propositions that are uncontroversial for *S* to assume tend to make *p* epistemically rational for him and no other propositions that are uncontroversial for him to assume can be used to defeat the support that these propositions give *p*. The idea is that a proposition *p* is epistemically rational for a person just if either *p* is properly basic for him or *p* is made epistemically rational for him by other propositions that are properly basic for him. (Foley, 1987:68)

Foley (1987) seems to be claiming that taking on a new belief is epistemically rational only if the new proposition, in virtue of the other beliefs he/she has, is uncontroversially true. To my mind this may be taking it too far, potentially ruling out the ability to take on unwanted beliefs or challenging beliefs. However, while Foley's claim may have different consequences to the model that I'm suggesting at least part of his claim appears to match on to the one above: namely that given the individuated Truth goal of each doxastic system only some new beliefs are rational to take on board. Others, in virtue of how they clash against the existing system of beliefs, are less likely to be believed. In these cases these propositions may require overwhelming

evidence to be believed by the agent, and, as I shall argue below, even this may not be enough in some cases.

In this section, I've argued that a doxastic system does not operate in isolation. This takes two forms. The first is that our other mental states can inform the way we form a specific belief. The second, the one that's germane to the claim that a doxastic system aims at rationality, is that our background beliefs, dispositions, emotional states, conations etc. shape the way our doxastic system goes about reaching the Truth goal.

5.3. Aiming at rationality and *idées fixe*

At the moment I feel as though my position is in a strong position to explain a number of features of belief in a comprehensive way.

That Truth is the primary goal of the system resolves the Problem of Motivation and explains why beliefs follow doxastic norms even unconsciously. A system behaving rationally would need to ensure that the cognitive processes respond to evidence. If they don't then they would no longer be fulfilling the goal of the system. Thus what accounts for the motivational power of epistemic reasons, causing beliefs to regulate for truth, is that the doxastic system as a whole is trying to rationally reach its goal.

But, the "shape" of each doxastic system (the other influencing factors) means that at times beliefs can come apart from Truth. In other words, sometimes preserving rationality entails or requires a failure to preserve Truth in order to meet the goal of tempered Truth. As a result, there are times when the beliefs we form are biased, or based on or sustained for non-epistemic reasons.

In addition, accepting that the doxastic system aims at rationality in the way that I've suggested above, explains why given the same evidence two different people might come up with two different beliefs. In both cases, the doxastic system is individually tailored by background cognitions, emotions, conations etc. This explains the scope of beliefs.

The question remains whether this account can make sense of *idées fixe*. In what follows below I argue that a system that aims at rationality is uncomplicatedly able to explain three of the types of *idée fixe* I laid out in Chapter 2. I leave delusions and why we're able to form beliefs on the basis of inconclusive evidence for later as part of a response to an objection to my account.

5.3.1. Bigoted belief and things I was going to believe anyway

In Chapter 2 I examined two different types of *idées fixe:* beliefs "... that I would … have no matter what philosophical arguments I was presented with" (Van Inwagen, 1993:15) and bigoted beliefs.

If a doxastic system aimed solely at fulfilling its goal of Truth both of these beliefs should suspend in the face of arguments to which I cannot give an epistemically satisfactory reply. That they don't suggests that my belief is being governed by a system that doesn't aim at Truth, *simpliciter*.

Now, if the goal of Truth is shaped in the way I've suggested above, then a doxastic system aiming at rationality would struggle to form beliefs that conflict too dramatically with our existing system of beliefs or our core beliefs. This is not to say that some event or evidence couldn't shift this belief, but that it seems some beliefs will require an awful lot of epistemic pushing to get them to budge.

What I'm suggesting is going on in the cases above is that the belief is entrenched. Giving up this belief, therefore, while it might be the right thing to do epistemically, is irrational given the agent's other (among others) beliefs.

Thus, while the belief may be false, it is, if we take the agent's full mental story into account, rational. What stops the person's belief engaging with the Truth is that the doxastic system is hindered from achieving Truth, *simpliciter*, by other commitments that a doxastic system needs to take into account in attempting to achieve its goal.

Van Inwagen's (1993) confession of the beliefs he was going to have anyway appears to be something that can be understood in terms of rationality, if not in terms of Truth.

Giving up believing something that is core to who you are as a person may be something that would call too many other beliefs into question; your whole system of truth would be undermined. Given your other commitments, therefore, and your beliefs at large, it would, in the sense suggested above, be irrational for you to give up this belief.

If something like this story is correct, then what makes it possible to be in a situation where there are "theses that I would believe no matter what philosophical arguments I was presented with" (Van Inwagen, 1993:15) is the fact that our doxastic system aims at rationality. If giving up one belief would completely unhinge a person's ability to make sense of the world, then a system that aims at preserving rationality may prevent one giving up that belief.[101]

The same story seems to be true of the bigot whose belief maintains in the face of counter evidence. Professor Shockley's racist beliefs upset and frustrate us. And they frustrate us not just because they're wrong, but because he seemingly refuses to give them up. As though this was something he could just "do" if he wanted to enough. Given the involuntarism of belief and a goal of Truth we ought to see his belief change in the face of the overwhelming counter evidence. So why don't we? My suggestion is that Shockley's belief informs so much of his life – his other beliefs, practices, academic esteem and so on – that to give up his belief, would in some sense be irrational. Giving up his belief could cause a doxastic crisis such that the complete shift in worldview would literally mean that it no longer makes sense. Thus, what explains the fact that his belief maintains in the face of counter evidence is the fact that giving it up would mean to give up a number of his other beliefs thereby undermining his ability to understand the world. A system that has the purpose of rationality would, therefore, be resistant to changing this belief.

[101] Talk of "preventing" isn't quite accurate actually. It's not that the doxastic system is engaging in something like self-deception – acknowledging that the Truth is but trying to preserve rationality to protect the agent – rather the Truth would not be believed because it wouldn't appear as Truth to the doxastic system, given the system's other commitments.

5.3.2. Beliefs that maintain in the face of unanswerable arguments

The most fascinating *idées fixe*, to my mind, are those beliefs that maintain in the face of unanswerable arguments. As I laid out in great detail in Chapter 2, Ribeiro (2002) asks why he can't believe the conclusion of epistemological scepticisms.

> I, BR, believe there *is* a cogent and intuitive argument for external world scepticism. I am quite convinced that this argument is unanswerable. … Strangely though, this argument does not move me in the least. (Though it is a strange personal confession to make) I have tried, quite seriously, to allow myself to be moved by it … But I am not moved. I still believe that I have a body and that my wife is real, that the floor before me really is there, and that I should act to insure the continued health of this apparent mortal coil. (Ribeiro 2002:475)

In this instance Ribeiro's (2002) belief in an external world remains, despite him acknowledging that he should, at the very least, suspend his belief in an external world given the evidence. I find Ribeiro's confession deeply troubling. Why don't we believe the conclusions to these types of arguments?

On the one hand I find arguments for external world scepticism enormously cogent and impossible to undermine. On the other, I find myself interacting with the world fully believing that it is real. As a result, I find G.E. Moore's (1939) common-sense response to external world scepticism as appealing as it is frustrating. It's appealing because the claim "Here's one hand, and here's another" (Moore, 1939:147) therefore an external world exists, ties in with my beliefs and provides me with a reason to hold my belief. However, it fails, at least to my mind, to even remotely adequately engage with strong arguments for external-world scepticism. So, how does a system that aims at rationality make sense of these beliefs?

I take it in these cases the belief the argument challenges is so central to the doxastic system as a whole that something quite remarkable would be needed to shift our belief. My belief in the external world is predicated on the fact that it (the external

world) makes itself known to me, sometimes quite painfully, on a daily basis. Similarly, when I think about the other beliefs I hold in the face of these unanswerable arguments (that I have free will and that I move), I find that the beliefs are so fundamental to what I believe about pretty much everything else that giving up these beliefs would, in some ways incapacitate me.

It doesn't hurt that, in terms of holding onto these beliefs, they constantly appear to be backed up by evidence every time I appear to move or appear to make a decision. Of course, the constant repetition, as I full well know, isn't sufficient evidence to determine Truth. Being told that 2+2=5 all the time doesn't make it True. However, given that these beliefs form part of the core of my doxastic system my claim is that a doxastic system operating properly would be shaped by this evidence in order to form rational, if not True, beliefs.

The reason that arguments against the existence of an external world don't work, I'm suggesting, is that they fail to shift the factors that shape the doxastic system. This is not to say that these arguments are in some way invalid, but to highlight that a stubbed toe or a hand in the fire provides a type of visceral consideration that René Descartes' (1996) arguments in favour of external world scepticism simply fail to emulate. Thus, regardless of whether our beliefs in the external world are True, they are rational.

In the case of each *idée fixe* above, therefore, what explains them is that they are rational. They actually meet the tempered goal of Truth. They are true from the agent's perspective.

Let's recap: a doxastic system that has Truth, *simpliciter*, as its goal could not account for *idées fixe*. In this case, the purpose of this system is met, if and when an agent actually believes the Truth. While this is certainly an attractive view, we could not explain evidence-resistant beliefs. If a doxastic system aims at rationality, however, it will govern beliefs such that they are motivated not only by the Truth goal, but also by the other beliefs, emotions, conations and so on an agent has. These tempering conditions shape the ability of the doxastic system to pursue Truth as the goal of the system and ultimately shape what truth is. Because of these tempering conditions, in

virtue of how the system is shaped, this system will paint a doxastic picture of the world that makes the most sense to the agent, all things considered.

In this case the all things considered are things like core beliefs, deep-seated feelings towards something, certain drives, desires, and so on. As a result, giving up an *idée fixe* would or could potentially throw out the whole system. This in turn could force the agent to find him/herself in a sort of doxastic crisis where the world stops making sense.

How would we expect a system that aims at doxastic rationality to react? It seems to me that what we could or should expect to happen is for the system to act in such a way as to preserve the rationality of the agent. To, if possible, ignore or refute the new evidence in order to preserve rationality. The result of this is, I suggest, evidence-resistant beliefs; beliefs that, if you will, have to maintain in order to preserve the rationality of agent.

Similarly, the homeostatic system responds to temperature fluctuations in the external environment in such a way as to preserve the core temperature of the agent. When an agent is exposed to higher than acceptable temperatures the agent sweats in order to try to preserve the correct core temperature. And in much the same way that we sweat in order to maintain homeostasis we maintain certain beliefs in order to preserve rationality. Evidence-resistant beliefs are, therefore, a sort of doxastic sweat; potentially smelly, but necessary to preserve the aim of the system in question.

This model, therefore, is able to account for three of the four *idées fixe* I discussed in Chapter 2. I have deliberately left out is how this model accounts for delusions. I examine delusions when I respond to one of the objections I level at my own account below. I turn now to these objections.

6. Critique

In this section I examine several possible problems with my account. I split these objections into two camps. The first are those objections which I believe my account can deal with. I deal with these in section 6.1. The second are those objections that I believe represent a real challenge to my account. I examine these in section 6.2. In 6.2 I suggest responses to these objections without really believing they are sufficient to overcome the challenge in question.

6.1. Objections I can deal with

The objections I want to deal with in this section are as follows:

1. A Tensionless Account equates Truth with truth (claiming that all of our beliefs are actually True). While other accounts rely on the gap between truth and Truth in order to create the necessary tension to explain the features of belief, a Tensionless Account does not have recourse to the same option. Therefore, a Tensionless Account suggests that all beliefs actually are True.
2. By setting up beliefs the way I have in Chapter 2 a Tensionless Account now begs the question. In Chapter 2 I argued that to believe that *p* is to take *p* as being True. Now, in Chapter 5 I've used that premise to prove the conclusion that the proper relationship between belief and Truth is a tensionless one. This begs the question.
3. Aiming at rationality isn't strong enough to account for beliefs "aiming at Truth". If a doxastic system takes on board other cognitions that shape this system – ultimately undermining its ability to reach its goal of Truth – how do we know that it doesn't sometimes fail altogether? Thus, while aiming at rationality may explain many of the potentially less palatable features of belief, it is unable to adequately capture the intuition behind Williams' (1973) notion that "beliefs aim at truth" (Williams 1973:136).
4. While we can't see it ourselves, it nevertheless seems to be the case that we hold beliefs that are contradictory. If a doxastic system aims at rationality in the way that I'm suggesting, this shouldn't be possible.

6.1.1. Conflating truth with Truth

The first objection to my account is that by suggesting that beliefs are in a tensionless relationship with Truth this means that I'm committed to claiming that every belief we have is in fact a correct belief.

Other accounts provide the explanatory tension required to explain why beliefs follow epistemic norms in virtue of the gap between belief and Truth. Thus, what does the work is the acknowledgement that there is a gap between what we believe (take as being True) and what is actually True.

I've noted that to believe that *p* is to take *p* as True, and that this is as far as the relationship between Truth and belief goes. The concern for my account is that given that I collapse the tension I may also be committed to the claim that what I take as True (my belief) actually is True.

However, while I deny that there is a tension, this does not imply that when we believe that *p* we believe the Truth; only that when we believe that *p* we take *p as* the Truth. Thus, when I believe that there is a computer in front of me I think that it is True that there is a computer in front of me. I could of course be wrong. It may be that I am in fact a brain in a vat being fed this sensory experience of sitting at a desk with a computer in front of me.

Denying this tension is to make the claim that the Truth doesn't provide any explanatory tension as it pertains to belief (irrespective of whether or not our belief actually is True). It is not to claim that what we believe is True in virtue of us believing it, only that we must think of our belief that it matches onto the Truth.

Thus, this account doesn't suggest a conflation between Truth and truth. All it acknowledges is that when an agent believes that *p* he/she thinks that he/she believes the Truth.

6.1.2. A Tensionless Account and begging the question

The second concern may be that by setting up the distinction between Truth and truth the way I have in the Chapter 2, and arguing that what an agent takes as True (the truth) is what a belief is, I'm now begging the question.

I'm using a premise (that to believe something is to take something as True) to prove my conclusion that a belief's relationship with truth is that when you take something as being True we call this state a belief.

I don't think this is a valid concern. All I'm saying is that the concept of belief is the full account of the explanatory tension Truth can provide with regards to belief.

6.1.3. Aiming at rationality isn't strong enough

The third concern is that a system that aims at rationality isn't able to solve the Problem of Motivation.

A system or thing that aims at Truth, *simpliciter*, can clearly explain the motivational power and authority of epistemic reasons. But can a system that "only" aims at rationality have enough power to explain both the authority and motivational force of epistemic norms?

What highlights a possible deficiency here is that I have postulated a system that has the aim of rationality in order to explain instances where beliefs are not only false, I've also argued that this system would allow for beliefs that are reluctant to regulate for truth. If this is correct, then how can we be sure that aiming at rationality is strong enough to account for the truth-tracking features of belief?

I argued above, that the system must, because it's a doxastic system, have Truth as its inherent purpose. But given that other states that shape the goal of Truth can take us away from Truth, how can we be sure that the Truth goal is strong enough to do the explanatory work it needs to in order to explain why our beliefs follow epistemic norms?

The first thing to note is that I'm not proposing an entirely closed system such that the system aims to maintain only a strict coherence between the existing set of beliefs. This proposal I suspect is ultimately self-defeating. Such a system could be accused, with some justification, of being unable to explain how we form new beliefs at all. The claim I'm making is that our existing set of beliefs inform the goal of Truth. While reaching Truth, *simpliciter*, may be subverted, the system, I'm arguing, is still trying to get things right.

Further, aiming at rationality in the way I've suggested will still have a clear link with Truth. While the other factors that shape the doxastic system may influence and occasionally jeopardise our ability to form beliefs in line solely with Truth considerations this it not to say that this rationality is not primarily epistemic in nature.

I've argued above that in some instances of *idées fixe* what explains their evidence-resistant nature is that the beliefs in question play such a vital role in the ability of the agent to function in the world that giving them up would be irrational. However, not all of our beliefs will be like this. In most cases a failure to take evidence on board would be equally irrational.

The agent who discovers that his/her car won't start would be irrational if he/she did not take this evidence on board and form a belief in line with this evidence. We can imagine that if he/she didn't this would dramatically undermine his/her ability to make sense of the world going forward.

Believing in line with the evidence is not only rational, in virtue of the goal it is also something that in itself aims at Truth. Aiming at rationality, therefore, requires aiming at Truth. If aiming at rationality requires us to take evidence into account then this more than adequately explains the authority and motivation that epistemic reasons have when it comes to our beliefs.

That said, it is a consequence of my model that the commitments that shape the goal of Truth could completely undermine an agent's ability to track the Truth. I take it

that this is precisely what happens in cases of delusions. In these cases, as a result of some trauma or some chemical imbalance, the goal of the doxastic system is completely undermined.

It is worth noting, though, that even in the extreme cases that delusions represent the doxastic system's goal of Truth can reassert itself. In a conversation I had with a psychologist colleague of mine she notes that delusions, such as the one laid out in Andrew Huddleston's (2012) paper in which "Joe" believes he is Napoleon, hardly ever maintain. The counter evidence is just too strong to ignore.

My model, therefore, appears to be able to do two things. The first is to provide an explanation for delusions. In these cases the other commitments I've spoken about radically impact on the doxastic system such that the beliefs deviate from what is taken as reality. The second is that even in the case of delusions we sometimes see the doxastic system reassert the goal of Truth.

Thus, while aiming at rationally fulfilling this tempered Truth goal can take us away from Truth it does seem to be strong enough to make sense of the Problem of Motivation.

Before I move on, it is also worth noting that my model of belief is able to explain why we are able to form beliefs on inconclusive evidence.

The model I'm presenting is of a doxastic system that aims at rationality as part of an agent's mental functioning as a whole. As a result, the aim to believe only the Truth, in line with the goal of the system, can be subverted by the other commitments of the agent.

Imagine, for example, that an evil scientist has captured your loved ones. She demands that you form a belief regarding the date of the first time man walked on the moon. She tells you that if you form the incorrect belief she will kill someone you love. In this instance, assuming you want your loved one to remain alive, you are likely to be meticulous in going about your evidence gathering for this belief. It would be rational, given what's at stake. Now, strictly speaking the instrumental costs are

not epistemic in nature and should not affect how a belief is formed. However, it seems silly to think that these instrumental concerns wouldn't play a role in this case of deliberation.

However, imagine you're at a dinner party and want to know if anyone can remember the date that Neil Armstrong first walked on the moon. In this instance when a member at the dinner party replies in full conviction that it was July 21, 1969, without even bothering to consult a reference, you might believe this without any further evidence. In this instance, given the context and your relative goals, this is perfectly rational.

What accounts for the differences in how we go about forming a belief are the other factors that shape the situation. What my account does is explain why it is that in some instances we don't engage in a rigorous evidence-gathering process in forming a belief.

Thus, not only is rationality strong enough to account for the truth-tracking features of belief it is also able to explain why this truth-tracking appears to be different in different contexts.

6.1.4. Contradictory beliefs

I take it as distinctly possible that we can have contradictory beliefs. While I don't think we can believe that *p* and at the same time and in full consciousness believe that not-*p*. Or believe that *p* and at the same time and in full consciousness believe that our belief that *p* is false. I do think that we can have contradictory beliefs.[102] To quote Michael Caine's character Nigel Powers in the spoof comedy *Goldmember* "There're only two things I hate in this world; people who are intolerant of other people's cultures and the Dutch". This quote suggests two contradictory beliefs: a belief that intolerance of other cultures is wrong; and a belief that the Dutch are people who one can be intolerant of.

[102] The former two claims are claims about what it is possible to believe about our beliefs, rather than a claim about whether it is possible for an agent to actually have contradictory beliefs. I take it that the former is impossible, despite what Andrew Huddleston (2012) says, while the latter is commonplace.

These beliefs contradict one another. If a system aims at rationality it is hard to see how this is possible. Holding onto contradictory beliefs is irrational, so how does a system that aims at rationality allow this to happen?

There are two potential responses here. The first, which has been suggested to me by David Martens, is as follows: what might be happening in this situation is the fact that while some beliefs appear to not fit into the rational schema, this might be because we're not looking broadly enough. A belief that *p* may be deemed as inconsistent with a certain subset of beliefs *q*. However, if we look at *p* as part of the system of beliefs as a whole or as part of the agent's mental functioning as a whole we will see that *p* is perfectly rational. What's going on when we say of someone else's belief that it is irrational, given their other beliefs, is that we're confusing the goal of the sub-schema (beliefs only) with either the doxastic system, which includes both beliefs as well as the non-belief sense-makers, or with the goals of the mental functioning of the agent as a whole.

The second, which ties into this suggestion, is that it is precisely because the system aims at rationality that an agent who holds contradictory beliefs cannot hold both of them at the same time in full consciousness. If Powers suddenly realised that one of his beliefs contradicted the other he would be forced to give up one of his beliefs. However, because he remains unaware of this he maintains both.

The irrationality would only come in when an agent sees that the two beliefs are mutually incompatible. However, both beliefs are important for Powers. They help him make sense of different situations. In this instance, a system that has the purpose of rationality would allow him to keep both.

6.2. Objections this account struggles to get around

In this section I examine two objections that potentially undermine my account.

1. Evidence-resistant beliefs can and do change.
2. Delusional beliefs don't even appear to aim at rationality.

These represent a serious challenge to my account. I explore these below providing a possible response where I have something.

6.2.1. Evidence-resistant beliefs can and do change

I argued above that in some cases a belief is so germane to how a person sees him/herself that giving up the belief would cause a massive disruption in the person's cognitive system, creating a doxastic crisis, potentially forcing an agent to question all of his/her beliefs.

This could, in extreme cases, incapacitate or seriously hinder someone's ability to interact with the world, meaning the world stops making sense. A system that aims at rationality, in virtue of aiming at rationality, protects the belief in order to make sense of the world; ultimately causing the belief to maintain in the face of counter evidence.

While this is a plausible explanation for evidence-resistant beliefs, and fits in with a doxastic system that aims at rationality, evidence-resistant beliefs can and do change. If the doxastic system aims at rationality, how is this possible? The only response I have to this is that in some instances it is perhaps even more irrational to continue to hold onto a belief in the face of overwhelming evidence to the contrary.

Consider the following: we have as much reason to believe that we live in the movie *The Matrix*'s matrix as we do to think that our world is the real world. However, despite this I certainly believe that I'm engaging with a real world and not a world that's been created for me by machines. The reason I hold onto this belief, despite the fact that the requirements of Truth suggest that I, at the very least, suspend my belief, is that it would incapacitate me to give up this belief. It would make me helpless. I couldn't be sure of any truth claim linked to this belief and so it would be epistemically irrational for me to give up this belief. Given that our doxastic system aims at rationality, my belief maintains in order to preserve this rationality.

However, if, perhaps, I was provided with overwhelming evidence to the contrary, then what I would find is my belief would change. Similarly, while the doxastic

system, in virtue of its goal of rationality, allows Shockley to maintain his racist belief by skewing the evidence he is presented with, perhaps if he was faced with incontrovertible evidence to the contrary his belief would change. Because at this point it would be less rational to hold onto his belief than it would be to give it up – despite the consequences giving up this belief would have.

While I think that this is part of the correct answer to this objection, I am not entirely happy with it.

6.2.3. Delusional beliefs don't even appear to aim at rationality

Part of what makes delusional beliefs delusional is precisely the fact that they are irrational. As Lisa Bortolotti and Matthew Broome (2008) note:

> Subjects with delusions are found to form their delusional state on the basis of evidence that is not strong enough to justify its wildly implausible content. Moreover, they typically hang on to their beliefs tenaciously, in a way that is impervious to what others take to be powerful counterevidence or counterarguments. The extent to which this occurs is so significant that it has been hypothesised that they suffer from a specific deficit which prevents them from rejecting the delusional belief as false, once it is formed. Finally, delusional states sometimes turn out to be badly integrated with other beliefs the subject has. (Bortolotti and Broome, 2008:822)

What this last sentence highlights is not only the failure of delusional beliefs to comply with the epistemic norms that govern the regulation of belief (something I think my model can make sense of) they "sometimes turn out to be badly integrated with other beliefs the subject has" (*ibid.*).

Huddleston (2012) makes a similar point in his example of Joe. He notes that not only is Joe's delusional belief evidence resistant it also conflicts with the other beliefs he has (Huddleston, 2012:214). A system that aims at rationality in the way that I'm

suggesting shouldn't allow for this any more than a system that aims at Truth, *simpliciter*, can account for this.

The temptation is to adopt one of two moves. The first, like Velleman (2000) and Robert Klee (2004), is to deny that delusions are beliefs at all. Velleman (2000) denies that delusions can be beliefs, because they are evidence resistant (Velleman 2000:277). Klee (2004), meanwhile, argues that "no one could literally believe what they say they do when they appear to be asserting a belief containing bizarrely delusional content" (Klee, 2004:27).[103] The second, is to adopt a move like the one that McKay and Dennett (2009) postulate above suggesting that the doxastic system has simply failed.

For me the first option is not an option at all. As I've argued in Chapter 2, because delusions display so many of the features of belief (including in rare instances regulating for truth) these should be properly understood as being delusional *beliefs*.

This leaves me with the second option, which is to claim that the doxastic system, in these cases, has simply broken down. This is enormously appealing. The fact that delusions are beliefs that are not "ordinarily accepted by other members of the person's culture or subculture" and maintain "despite what constitutes incontrovertible and obvious proof or evidence to the contrary" (Klee, 2004:25) certainly could suggest that something has gone wrong. Moreover, these beliefs that are at odds with the other members of the person's society are not like believing in racial equality during the *apartheid* regime in South Africa. They are genuinely odd. People who suffer from the Cotard delusion, for example, believe they are dead (Klee, 2004:26). In addition, the fact that sometimes these beliefs aren't even internally

[103] Robert Klee (2004) acknowledges that not all delusions are internally irrational. As he puts it:

> … often a deluded person acts in rational on the basis of the apparent content of a delusion, a delusion that is itself irrational given external rationality. Ironically we render the sufferer's behaviour rational by attributing to him or her a belief that is itself irrational. (Klee, 2004:27)

If all delusions were like this, then I don't think my account would have a problem. In fact, my account would appear to be bolstered by this. However, the key part of the quote above is that these people are "often", but not always, rational. Thus, if there are even some instances in which the delusions are irrational, my account is in trouble.

rational indicates that something isn't quite right. And, finally, those who suffer from delusions are routinely treated in a way that suggests a disease or misfunctioning.

However, despite the appeal I'm not happy with any of these reasons. The fact that our beliefs don't cohere with the majority is not in and of itself a reason to suggest that the mental state is a result of a broken mechanism. Further, pointing to the oddity of the belief in question appears to be a matter of degree, rather than type. If this is correct, then a proper account of the truth-belief relationship ought to be able to capture these beliefs.

The second reason, namely that these beliefs aren't internally rational and therefore are the result of a broken doxastic system, seems question begging. What I'm trying to explain, using a system that as part of its functioning aims at this kind of rationality, are delusional beliefs. That there are some beliefs that don't fit into this system suggests there's something wrong with the system or with the belief. But to claim that it's not the system, but the belief is at best table-thumping and at worst question begging.

Finally, that we tend to medicate or lock up people with delusions is not a reason to think that these beliefs are the result of a broken doxastic system. All it is is an acknowledgement of how we as a society tend to treat people whose beliefs fall outside what we consider an acceptable norm.

7. Conclusion

While delusions – or some delusional beliefs – represent a significant challenge to my account I think that there are nevertheless a number of benefits to my account.

I've argued that other accounts fail for a host of reasons. I have argued that the primary flaw of these other accounts is the fact that they postulate a dynamic relationship between beliefs themselves and Truth in order to make sense of the features of belief. This tension ultimately undermines these accounts of the truth-belief relationship.

I've argued that removing the tension between belief itself and Truth explains several features of belief and, importantly, allows us to keep intact what it means to be a belief. What provides the tension required to explain why beliefs follow epistemic norms, I've argued, is a doxastic system of which beliefs are just one part.

This move, I've argued, solves several problems simultaneously. First of all it accounts for the Problem of Motivation by proposing a constitutive goal of the doxastic system. Second of all, it allows us to hold onto the Tensionless Account with all its benefits. These benefits include the ability to explain several features of belief as well as the fact that we no longer require the belief *qua* belief to make sense of, among others, the Problem of Motivation, which allows us to place belief and truth in the correct relationship with one another.

I argued that this doxastic system has a goal of Truth, which accounts for the Problem of Motivation. However, this goal of Truth is tempered by an agent's other commitments and so the system *de facto* aims at rationality. A system that aims at rationality, I argued, requires that a potential belief fits in with his/her other experiences, beliefs, perceptions and so on. All of these are Truth-bearing and along with his/her other beliefs, background or foreground, would contribute towards how he/she understands the proposition under consideration (whether that consideration is done in full consciousness, as is the case of deliberation, or unconsciously).

This system makes sense of not only the Problem of Motivation, but in virtue of these tempering factors can explain why we form beliefs on inconclusive evidence and *idées fixe*.

That said, it is not without its problems. Delusions of a certain kind represent a real challenge to my system as do *idées fixe* that ultimately regulate for truth.

Chapter 6: Conclusion

The aim of my thesis has been to try to find a model of the truth-belief relationship that is able to explain several features of belief in virtue of the suggested relationship between beliefs and truth.

I examined two models of the truth-belief relationship in the extant literature: teleological accounts and normative accounts. I argued that neither of these models sufficiently captured the relationship between beliefs and truth in the way that I was exploring. I then suggested a third model, which I've argued has far more explanatory power in terms of the features of belief than the other two models.

In Chapter 2, I laid out those features of belief that I think an account of the truth-belief relationship needs to account for. In addition, I made clear an argument in the literature that it is the job of the truth-belief relationship to account for the features of belief.

In Chapter 3 I looked at teleological accounts' ability to explain the features of belief. I argued that teleological accounts, both the exposition and the criticism of, are made more complicated by the fact that there is a distinction missing from the literature.

I argued that teleological accounts can be understood in two distinct ways: either as aims or as purposes. In the former, the belief would aim at the truth in much the same way that archers aim at targets, while in the latter beliefs would have truth as their purpose in much the same way that boiling water is the purpose of kettles.

Despite this distinction teleological accounts, whether as Aim Accounts or Purpose Accounts fail to make sense of the truth-belief relationship. In both cases, these accounts fail to make sense of several key features of belief; the most damning of which is the failure to keep the link between beliefs and truth direct and the inability to explain *idées fixe*.

In Chapter 4 I explored the ability of Normative Accounts to make sense of the truth-belief relationship. To begin with, I noted that as it stands the most widely-cited

examples of a Normative Account of the truth-belief relationship, those of Nishi Shah (2003) and later Shah and David Velleman (2005), are in fact examples of a deliberation-truth relationship. Because of this, I argued, they cannot be seen as legitimate accounts of the truth-belief relationship. Having made this argument I developed a Normative Account using their arguments as a base.

This model, however, as with teleological accounts, fails as an adequate account of the truth-belief relationship. What undermines these accounts, I argued, was the fact that they cannot account for why we form beliefs on the basis of inconclusive evidence, nor why some of our beliefs are evidence resistant.

In Chapter 5 I presented my own model of the truth-belief relationship. This model of the truth-belief relationship has two parts. The first argues in favour of a non-dynamic relationship between beliefs and truth. I argued in this chapter that by placing Truth and belief in a Tensionless Account helps us avoid the problems that are inherent in Aim Accounts and Purpose Accounts. I argued, therefore, that the right way to understand the relationship between a belief itself and Truth is as a tensionless relationship. That the full extent of the tension between beliefs and Truth is that to believe that *p* is to take *p* as being True.

This model explains a number of features of belief in an unproblematic manner. In particular it makes sense of those features that indicate the direct link between beliefs and truth.

However, this Tensionless Account, while able to explain a number of features of belief cannot, in and of itself, explain "the authority of epistemic norms but also their motivational efficacy" (Owens 2003:284). I called this the Problem of Motivation.

I argued that if beliefs are part of a doxastic system (controlling the formation, revision and retention of beliefs) that has the purpose of Truth we can account for the Problem of Motivation.

In much the same way that teleological accounts can appeal to the constitutive goal of belief in order to make sense of the Problem of Motivation, the suggestion here was

that a doxastic system can appeal to the constitutive goal of the system to make sense of the Problem of Motivation.

However, I argued that this goal of Truth should be seen as a tempered goal of Truth. Other cognitions, conations and emotions shape the goal of Truth such that we can meaningfully talk about a doxastic system that aim, *de facto*, at rationality. What made this move appealing was that it was able to make sense of the type of interplay beliefs have with other psychologically motivating, though epistemically clouding, mental states and attitudes.

This acknowledgement of the impact that other mental states have on our ability to pursue the goal of Truth, *simpliciter*, I argued, accounts for, among others, epistemic bias, wishful thinking and *idées fixe*.

My account, that beliefs are part of a doxastic system that aims at rationality, I argued was able to make sense of several of the *idées fixe* that I've explored throughout this dissertation.

I think one of the real advantages of this position is that it helps explain why sceptical arguments – the problem of induction, external world scepticism and so on – are so hard to actually believe. For many, David Hume's (1977) problem of induction represents a clear case of an argument that is both unanswerable and unbelievable. Hume makes it quite clear that to form a belief based on the process of induction is not a very good way of getting to the truth. However, despite this clear and cogent argument I certainly find it hard to believe. I form beliefs on the basis of induction all the time. Similarly, I certainly find René Descartes' (1996) arguments in favour of external world scepticism impossible to undermine to the satisfaction of my philosophical and epistemic desire, and yet, I find it impossible to believe that there is no external world.

If you'll forgive the brief digression, I have often wondered, quite seriously, why philosophers play pool, or don't continuously walk into walls. If we take Hume's (1977) argument against induction seriously, playing pool is at best a gamble and at worst entirely pointless. There is after all, given the validity of Hume's argument, no

guarantee that hitting a ball in a certain direction with a certain force will cause the ball to go in the direction that you want it to. Similarly, if we take some argument for external-world scepticism seriously, and again as philosophers I feel we ought to, it seems perfectly epistemically valid to attempt to try to walk through a wall. Given the arguments for external world scepticism, postulating the existence of walls is horribly unjustified epistemically speaking.

Yet, despite the epistemic seriousness of both of these arguments we don't see philosophers walking into walls (well, not often, and never, in my experience, because he/she didn't believe the wall was there as a result of some sceptical argument) and the philosophers I have met appear to have a great penchant for pool, if not snooker.

The fact that we as philosophers who, in theory, take reason very seriously, don't believe in line with these arguments has baffled me.

A doxastic system that aims at rationality, I hope, to my lasting relief, explains why this is the case. The pool-playing philosopher believes that playing pool is a legitimate and efficacious way to spend his/her leisure time. Induction may be a bad way of getting to the truth, but it is remarkably persistent. Given a system that aims at rationality, induction, given that it is the basis for so many of our beliefs, would be something that it would epistemically rational to believe in.

Further, discarding induction as a method of reasoning, it seems, would, it seems create an incredibly uncertain and scary world. Consider what it would be like to genuinely believe that the sun might/could/would not rise tomorrow; that your office was potentially somewhere else; or that the way you went to your office yesterday was not a reliable way of getting there today. Much of the world as we know it (and I use the word "know" quite deliberately) would stop making sense.

A system that aims at the Truth, *simpliciter*, however, it seems, is committed to putting you into precisely this position by suspending your belief in the epistemic worth of induction in favour of waiting for the Truth. You could imagine a doxastic system that aims at Truth commenting: "look, you can try induction if you want to, but I wouldn't have any beliefs about it if I were you". While a system that aimed at

Truth, *simpliciter*, might be epistemically appealing I question whether this system would allow us to actually cope in the world.

Hume (1977), himself, thought we needed to believe in the epistemic worth of induction, despite the fact that he felt there was no good reason to trust this method of reasoning. His conclusion was that we continue to believe that induction gets to the Truth as a result of "Custom or Habit" (Hume 1977:28).

However, this answer fails to adequately explain *why* custom or habit maintains this belief in the face of counter evidence when other beliefs respond quite readily to counter evidence.

The answer, I suggested, is that it would be irrational to believe that induction fails; irrational in the sense that any person who believed Hume's (1977) sceptical argument would live in a world that stopped, for the most part, making sense.

Thus, given a system that aims at rationality we can see not only why we believe the conclusions of induction, but why we believe that playing pool is a worthwhile (if sometimes incredibly frustrating) past-time.

Similarly, once you've walked into a few walls it is rational, if not perhaps correct, to believe that the next wall you encounter will be an unyielding and painful barrier to your perambulation between areas. Thus, what makes these scepticisms, dare I say it, impossible to believe is the fact that it would be irrational to believe them. An actual belief in any of them would put us into real trouble, very quickly.

What explains why we hold onto these beliefs is that we have a doxastic system that aims at rationality. In doing so it causes these beliefs to maintain despite strong arguments to the contrary because given all the other commitments of the system it is true, if not True, that the external world exists and that induction is a reliable method of reasoning.

Clark, A. (2005). "Thought in a Hostile World: The Evolution of Human Cognition by Kim Sterelny", *Mind*, 114(455), pp 777-782

Clifford, W.K. (1999). "The Ethics of Belief". In Joel Feinberg and Russ Shafer-Landau (ed.), *Reason and Responsibility: Readings in Some Basic Problems of Philosophy* (10th Edition). Belmont CA: Wadsworth. Pp 121-125

Corvino, J (1997). "Why Shouldn't Tommy and Jim Have Sex? A Defense of Homosexuality", In John Corvino (ed.) *Same sex: debating the ethics, science, and culture of homosexuality*, ed. John Corvino. Lanham, Md: Rowman & Littlefield

Cottingham, J. (2002). "Descartes and the Voluntariness of Belief", *The Monist*, 85 (3), pp 343-360

Davidson, D. (1986). "A Coherence Theory of Truth and Knowledge". In Ernest LePore (ed.), *Truth and Interpretation*. Cambridge: Blackwell. Pp 307-319

De Cruz, H., Boudry, M., De Smedt, J. & Blancke, S. (2011). "Evolutionary Approaches to Epistemic Justification", *Dialectica*, 65 (4), pp 517-535

Della Rocca, M. (1996). "Recent Work: Essentialism: Part 1", *Philosophical Books*, 37, pp 1–13

Della Rocca, M. (1996a). "Recent Work: Essentialism: Part 2", *Philosophical Books*, 37, pp 81–89

Dennett, D. (1987). *The Intentional Stance*. Massachusetts: Massachusetts Institute of Technology

Dennett, D. (2004). "An evolutionary perspective on cognition: through a glass lightly", *Studies in History and Philosophy of Biological and Biomedical Sciences*, 35, pp 721-727

Descartes, R. Ed. Cottingham, J. (1996). *Meditations on First Philosophy*. Cambridge: CUP

Diamond, C. (2008). "Anything but argument?", *Philosophical Investigations*, 5 (1), pp 23-41

Edmonds, A.K. (2007). *Abandoning The Truth Aim A Reevaluation Of The Aim Of Belief And The Goal Of Cognition*. (Unpublished Honors thesis, Williams College)

Feldman, R. (2000). "The Ethics of Belief", *Philosophy and Phenomenological Research*, 60, pp 667-695

Fine, K. (1994). "Essence and Modality: The Second Philosophical Perspectives Lecture", *Philosophical Perspectives*, 8, pp. 1-16

Fine, K. (1994a). “Senses of Essence” In Walter Sinnot-Armstrong (ed.), *Modality, morality and belief Essays in honor of Ruth Barcan Marcus*. Cambridge: Cambridge University Press. Pp 53-73

Frankfurt, F. (1969). “Alternate Possibilities and Moral Responsibility”, *Journal of Philosophy*, 66, pp 828-839

Gendler, T.S. (2008). “Alief and Belief”, *Journal of Philosophy*, 105 (10), pp 634-663

Gettier, E.L. (1963). “Is Justified True Belief Knowledge?”, *Analysis*, 23 (6), pp 121-123

Gibbard, A. (2005). “Truth And Correct Belief”, *Philosophical Issues*, 15, pp 338-350

Ginet, C. (2001). “Deciding to Believe”. In Matthias Steup (ed.), *Knowledge, Truth and Duty*. Oxford: Oxford UP. Pp 63-76

Glanzberg, M. (2009). “Truth”, *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2009/entries/truth/>

Heller, M. (2000). “Hobartian Voluntarism: Grounding a Deontological Conception of Epistemic Justification”, *Pacific Philosophical Quarterly*, 81 (2), pp 130-141

Heyes, C.M. & Huber, L. (Eds.) (2000). *The Evolution of Cognition*. MIT Press: Cambridge

Hieronymi, P. (2006). “Controlling Attitudes”, *Pacific Philosophical Quarterly*, 87, pp 45-74

Huddleston, A. (2012). “Naughty beliefs”, *Philosophical Studies*, 160, pp 209-222

Huggett, N. (2010). “Zeno’s Paradoxes”, *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2010/entries/paradox-zeno/>

Humberstone, I.L. (1992). “Direction of Fit”, *Mind*, 101 (401), pp 59-83

Hume, D. Ed. Selby-Bigge, L.A. (1978). *A Treatise of Human Nature.* Oxford: Clarendon Press

Hume, D. Ed. Steinberg, E. (1977). *An Enquiry Concerning Human Understanding*. Cambridge: Hackett Publishing Company Indianapolis

James, W. (1999). In Joel Feinberg and Russ Shafer-Landau (ed.), *Reason and Responsibility: Readings in Some Basic Problems of Philosophy* (10th Edition). Belmont CA: Wadsworth. Pp 125-133

Johnston, M. (1988). “Self-Deception and the Nature of Mind”. In B.P. McLaughlin and A.O. Rorty (eds.), *Perspectives on Self-Deception.* University of California Press: Berkeley. Pp. 63–91

Jones, W.E. (2002). “Explaining Our Own Beliefs: Non-epistemic Believing And Doxastic Instability”, *Philosophical Studies*, 111 (3), pp 217-249

Jones, W.E. (2004). “Pragmatic Believing And Its Explanation”, *Critica*, 36 (108), pp 3-36

Jordaan, E. (2012). “Schizophrenia”. In A. Burke (ed.), *Abnormal Psychology: A South African Perspective* (2nd Edition). Southern Africa: Oxford University Press. Pp. 190-237

Kavka, G.S. (1983). “The Toxin Puzzle”, *Analysis*, 43 (1), pp 33-36

Kelly, T. (2003). “Epistemic Rationality as Instrumental Rationality: A Critique”, *Philosophy and Phenomenological Research*, 66, 612-640

Klee, R. (2004). “Why Some Delusions Are Necessarily Inexplicable Beliefs”, *Philosophy, Psychiatry & Psychology*, 11 (1), pp 25-34

Klement, K. (2011). “Russell’s Logical Atomism”, *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2011/entries/logical-atomism/>.

McHugh, C. (2012). “Belief and aims”, *Philosophical Studies*, 160 (3), pp 425-439

McKay, R.T. & Dennett, D.C. (2009). “The evolution of misbelief”, *Behavioural and Brain Sciences*, 32, pp 493-561

Montmarquet, J. (1986). “The Voluntariness of Belief”, *Analysis*, 46 (1), pp 49-53

Moore, G.E. (1939). “Proof of an external world” *Philosophical Papers* (Vol. 25). New York: Collier Books. Pp 126-148

Nasar, S. (March 25, 2002). “The sum of a man”. *The Guardian*. Accessed 29 November 2013

Nasar, S. (November 13, 1994). “The Lost Years of a Nobel Laureate”. *The New York Times*. Accessed 29 November 2013

Nisbet, E. & Garrett, K. (October 27, 2010). “Belief In Rumors Hard To Dispel: Fact checking easily undermined by images, unrelated facts” URL = <http://www.comm.ohio-state.edu/kgarrett/FactcheckMosqueRumors.pdf>

Owens, D.J. (2002). “Epistemic Akrasia”, *The Monist*, 85(3), pp 381-397

Owens, D.J. (2003). “Does Belief Have an Aim?”, *Philosophical Studies*, 115, pp 283-305

Papineau, D. (2004). “Friendly Thoughts on the Evolution of Human Cognition”, *Australasian Journal of Philosophy*, 82 (3), pp 491-502

Pascal, B. Ed. Ariew, R. (2005). *Pensees*. Indianapolis: Hackett

Plato. (2002). *The Republic*. Available from: http://www.idph.net. (Accessed 1 July 2012)

Quine, W.V. & Ullian, J.S. (1978). *The Web of Belief* (2nd Edition). New York: Random House

Rachels, J., & Rachels, S. (2010). *The Elements of Moral Philosophy* (6th Edition). New York: McGraw-Hill

Ribeiro, B. (2002). "Epistemological Skepticism(s) and Rational Self-Control" *The Monist*, 88 (3), pp 468-477

Rodman, F.R. (2003). *Winnicott*. Cambridge: Perseus Books

Rosati, C.S., (2008). "Moral Motivation", *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2008/entries/moral-motivation/>

Samovar, L.A., Porter R.E., & McDaniel, E.R. (2010). *Communications between cultures* (7th Edition). Boston: Centrage

Sartre, J.P. (1956). *Being and Nothingness*. USA: Washington Square Press

Schwitzgebel, E. (2010). "Belief", *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2010/entries/belief/>

Scott-Kakures, D. (1994). "On Belief and Captivity of the Will", *Philosophy and Phenomenological Research*, 54, pp 77-103

Scott-Kakures, D. (2000). "Motivated Believing: Wishful and Unwelcome", *Noûs*, 34 (3), pp 348-375

Shah, N. & Velleman, D.J. (2005). "Doxastic Deliberation", *The Philosophical Review*, 114 (4), pp 497-534

Shah, N. (2002). "Clearing Space for Doxastic Voluntarism" *The Monist*, 85 (3), pp 436-445

Shah, N. (2003). "How Truth Governs Belief", *The Philosophical Review*, 112 (4), pp 447-482

Steglich-Petersen, A. (2006). "No Norm Needed: On The Aim Of Belief", *The Philosophical Quarterly*, 56 (225), pp 499-516

Steglich-Petersen, A. (2009). "Weighing the aim of belief", *Philosophical Studies*, 145, pp 395-405

Stephens, C. (2001). "When is it Selectively Advantageous to have True Beliefs? Sandwiching the Better Safe than Sorry Argument", *Philosophical Studies*, 105 (2), pp 161-189

Sterelny, K. (2003). *Thought in a Hostile World: The Evolution of Human Cognition*. Oxford: Blackwell Publishing

Steup, M. (2000). "Doxastic Voluntarism and Epistemic Deontology", *Acta Analytica*, 15 (24), pp 25-56

Stoneham, T. (1998). "On Believing That I Am Thinking", *Proceedings of the Aristotelian Society*, New Series, 98, pp 125-144

Stroud, S. (2006). "Epistemic Partiality in Friendship", *Ethics*, 116, pp 498-524

Swartz, L., De La Rey, C. & Duncan, N. (2004). *Psychology an introduction*. New York: OUP

Vahid, H. (2006). "Aiming At Truth: Doxastic vs Epistemic Goals", *Philosophical Studies*, 131, pp 303-335

Van Fraassen, B.C. (1984). "Belief and the Will", *Journal of Philosophy*, pp 235-256

Van Inwagen, P. (1993). *Metaphysics*. New York: OUP

Van Inwagen, P. (2009). *Metaphysics* (3rd Edition). New York: OUP

Velleman, J.D. (2000). *The Possibility of Practical Reason*. New York: OUP

Viedge, N. (2004). *Knowing What We Can't Believe*. (Unpublished MA dissertation, Rhodes University)

Wedgwood, R. (2002). "The Aim of Belief", *Philosophical Perspectives*, 16, pp 267-297

Whiting, D. (2013). "Nothing but the Truth: On the Norms and Aims of Belief". In Timothy Chan (ed.), *The Aim of Belief*. Oxford: Oxford University Press.

Williams, B. (1973). "Deciding to Believe" In Bernard Williams (ed.), *Problems of the Self*. Cambridge: CUP. Pp 136-151

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

Winters, B. (1979). "Believing at Will" *Journal of Philosophy*, 76 (5), pp 243-256

Wiredu, K. (1980). *Philosophy and an African Culture*. Cambridge: Cambridge University Press