THE TRUTH NORM ACCOUNT OF JUSTIFICATION

Alexander Greenberg
Trinity Hall
University of Cambridge

6th July 2017

This dissertation is submitted for the degree of Doctor of Philosophy
THE TRUTH NORM ACCOUNT OF JUSTIFICATION

ABSTRACT

This thesis is about the relationship between a belief being justified and it being true. It defends a version of the view that the fundamental point of having a justified belief is to have a true one. The particular version of that view it defends is the claim that belief is subject to a truth norm – i.e. a norm or standard that says that one should believe something if and only if it’s true. It claims that belief being subject to such a truth norm can explain which beliefs count as justified and which do not.

After introducing the idea of a truth norm (Ch. 1), the argument of my thesis involves two main stages.

Part One of the thesis (Chs. 2-3) contains the first stage, in which I argue that my way of arguing for a truth norm, on the basis of its explanatory role in epistemology, is much more likely to be successful than a more popular way of arguing for a truth norm, on the basis of its explanatory role in the philosophy of mind.

Part Two (Chs. 4-7) contains the second stage, in which I argue that the truth norm can indeed explain justification in the way I’ve outlined. I do this by answering four criticisms that have been made of the claim that belief is subject to a truth norm. These criticisms claim that a truth norm should be rejected because, in turn, a truth norm cannot guide belief formation (Ch. 4), because a truth norm prescribes believing all the truths (Ch. 5), because a truth norm never prescribes suspending judgement (Ch. 6), and because a truth norm in some cases prescribes making problematic trade-offs of having one false belief for the sake of having many true beliefs (Ch. 7). I argue that all of these criticisms fail. But it is through answering these criticisms that we can see the contours of a defensible explanation of justification in terms of the truth norm.

Alexander Greenberg
July 2017
DECLARATION

This thesis is 78,612 words in length. The word count, including footnotes and references, falls within the range specified by the Degree Committee of the Faculty of Philosophy. This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University.
ACKNOWLEDGEMENTS

For financial support, I’d like to thank the Arts and Humanities Research Council for funding this thesis, the Cambridge Philosophy Faculty and an anonymous donor for writing-up grants, and my college, Trinity Hall, for travel grants.

I’d like to thank my supervisors, Tim Crane and Arif Ahmed. Throughout this thesis I owe so much to their help. I’d also like to thank Tim in particular for continuing to support and to have faith in me during the times in which work was difficult, especially in the periods before and after my intermission. I also thank my examiners, Pascal Engel and Richard Holton, for a rich and helpful discussion of the thesis.

Special thanks are due to Lucy Campbell and to Chris Cowie for countless discussions, which aided me immeasurably in writing this thesis. For various philosophical discussions over the past five years, I also thank Lukas Skiba, Dan Brigham, Fiona Doherty, Nakul Krishna, Bernard Salow, Lubomira Radoilska, Louise Hanson, Rob Trueman, Tim Button, Rae Langton, Jane Heal, and Hugh Mellor. Thanks are in particular due to those who have read and commented on work that would form part of this thesis, Lucy Campbell, Chris Cowie, Nakul Krishna, Fiona Doherty, Bernard Salow, Selim Berker, and Conor McHugh. I’m also grateful to audiences in Cambridge, Edinburgh, Fortaleza, and King’s College London, who listened to and commented on previous incarnations of various parts of this thesis.

I’d also like to thank the various friends, both old and new, who have kept my spirits up over the last five years, in particular Lucy Campbell, Chris Cowie, Lukas Skiba, Ellisif Wasmuth, Jenni Sidey, Nakul Krishna, Fiona Doherty, Dan Brigham, Christopher Clarke, Ian Felce, Raphael von Blumenthal, Mark Johnson, Jack Miller, Sam Baron, James Dowman, Simon Coplowe, and Ross Frame. I’d also like to thank the members of the Trinity Hall rugby team, and the gardens teams of Trinity Hall and King’s College Cambridge (especially Helen Cripps), with whom I worked during my intermission. I’d also like to thank Chris Rowland and Heather Sanderson.

Special thanks are due to my family – my parents in particular – for putting me up (and putting up with me) for a year when I had nowhere else to live. Thanks to my erstwhile roommate Eggbert for keeping me company. And most of all thanks to Lucy Campbell for keeping me sane, and for tirelessly reminding me, despite my protestations, that writing this thesis was possible.
# Table of Contents

Summary of Chapters........................................................................................................1

List of Norms and Principles.............................................................................................5

Chapter 1. Introduction........................................................................................................9
1. The truth norm account of justification in context .......................................................10
1.1. Veritism ....................................................................................................................10
1.2. Justification as a normative notion ............................................................................12
2. How a truth norm explains justification .......................................................................16
2.1. Objective and subjective norms on belief .................................................................19
2.2. What are the ‘good means’ of conforming to a truth norm? .................................22
3. Conclusion..................................................................................................................26

Part Two. How and How Not to Argue for a Truth Norm .................................................29

Chapter 2. How Not to Argue For a Truth Norm ..............................................................31
1. Can a truth norm explain what distinguishes belief from other attitudes? ...............32
2. Can a truth norm explain the normativity of content? ...............................................37
3. Can a truth norm explain the transparency of doxastic deliberation? ....................40
4. Can a truth norm explain the impossibility of believing at will? ............................47
5. Can a truth norm explain Moore’s paradox? .............................................................50
6. A better way: belief is subject to a truth norm because that explains epistemic justification.................................................................................................................54
6.1. The appeal to this methodology in ethics...............................................................57
6.2. A limitation of this methodology ...........................................................................62
7. Conclusion..................................................................................................................63

Chapter 3. The Unity of Justification..................................................................................65
1. Different methods that justify beliefs ..........................................................................65
1.1. Why something must unify different methods that justify beliefs.........................66
1.2. What could unify different methods that justify beliefs? .......................................69
1.3. How a truth norm unifies different methods that justify beliefs .........................72
2. Different properties that bear on the justification of a belief ....................................75
Chapter 6. Suspension of Judgement

1. The suspension-of-judgement problem

2. Engel's appeal to evidential norms

2.1. What it is to have ‘sufficient’ evidence

2.2. Why following MORE THAN 50-50 is a better means of conforming to OUGHT

3. The solution: give greater weight to the aim of avoiding error

3.1. Wedgwood: an intermediate degree of correctness

3.2. My solution: a contrary-to-duty obligation to err on the side of caution

4. Conclusion

Chapter 7. Sacrifices for the Greater Epistemic Good?

1. Veritism and Firth cases

2. An answer: proposition-relative veritism

3. Objections and replies

3.1. The problem of weighing proposition-relative goals against each other

3.2. The problem of self-fulfilling beliefs?

4. Conclusion

Conclusion

Bibliography
SUMMARY OF CHAPTERS

This thesis is about the relationship between a belief being justified and it being true. It defends a version of the view that the fundamental point of having a justified belief is to have a true one. The particular version of that view it defends is the claim that belief is subject to a truth norm — i.e. a norm or standard that says that one should believe something if and only if it’s true. It claims that belief being subject to such a truth norm can explain which beliefs count as justified and which do not.

After introducing the idea of a truth norm in Chapter 1, the argument of my thesis involves two main stages. Part One of the thesis (Chs. 2-3) contains the first stage, in which I argue that my way of arguing for a truth norm, on the basis of its explanatory role in epistemology, is much more likely to be successful than a more popular way of arguing for a truth norm, on the basis of its explanatory role in the philosophy of mind. Part Two (Chs. 4-7) contains the second stage, in which I argue that a truth norm can indeed explain justification in the way I’ve outlined.

PART ONE: HOW AND HOW NOT TO ARGUE FOR A TRUTH NORM

In Chapter 2, I discuss various attempts to argue that belief is subject to a truth norm on the basis of the explanatory role of such a norm in the philosophy of mind. Such attempts typically argue that we need to think of the attitude of belief as constitutively subject to a truth norm in order to explain some psychological feature of belief. It has been suggested that belief being subject to a truth norm can explain, for example, what distinguishes belief from other attitudes, the transparency of doxastic deliberation, the normativity of mental content, the impossibility of believing at will, and Moore’s paradox. These attempts are unconvincing, I argue, because either these features can be explained by much weaker claims than a truth norm, or a truth norm can only explain these features given very strong theoretical assumptions.

I conclude that we should not try to give independent reason for thinking belief is subject to a truth norm, a norm which we then go on to use to explain justification. Instead, I argue that we should think of belief as subject to a truth norm because that can do justice to our pre-theoretical understanding of justification, much as one might argue for a moral norm or theory because it does justice to our pre-theoretical understanding of what we ought to do. The best argument for a truth norm is not one that appeals to philosophy of mind, but one
within epistemology.

Chapter 3 goes into detail about a key feature of our pre-theoretical understanding of justification which a truth norm can explain. There I argue that a truth norm can explain the unity of justification, i.e. what unifies justified beliefs that in other respects look heterogeneous. In particular, I argue that a truth norm can explain a) what unifies radically different methods that lead to justified beliefs (e.g. what explains why both mental arithmetic and testimony can lead to justified beliefs); and b) what unifies the radically different properties of a hypothesis that bear on whether one is justified in believing it (e.g. how conservative, modest, or simple the hypothesis is).

**PART TWO: HOW A TRUTH NORM EXPLAINS JUSTIFICATION**

In Part Two, I argue that a truth norm can explain justification. I do this by answering four criticisms of such a norm. Responding to these criticisms involves modifying and amending the basic truth norm account in a few different ways. We will see that once it is modified, a truth norm can succeed in explaining justification. Thus it is by answering these criticisms that we will have in view the contours of a defensible explanation of justification in terms of a truth norm.

Chapter 4 discusses an argument against a truth norm account given by Katrin Glüer and Åsa Wikforss, which is known as ‘the no guidance argument’. This argument claims that we should reject a truth norm because it cannot guide belief formation, because in order to follow it one has to determine whether p is true, which amounts to forming a belief about p. This makes the guidance seem to come ‘too late’. I show that their argument fails. I first argue that it relies on an overly restrictive conception of what it is for a norm to provide guidance, drawing on a response given by Asbjørn Steglich-Petersen. We should instead prefer a more inclusive conception of guidance. Secondly, given that the central role of a truth norm is to explain why one should follow other norms, a truth norm does not need to actively provide guidance in a believer’s consciousness in order to play this explanatory role.

Chapter 5 answers philosophers who object to truth norms which prescribe believing all the truths. The way most proponents of a truth norm respond to this problem is to put forward a modified truth norm that doesn’t make this implausible prescription. I show that a number of these modifications have been missteps. These include making a truth norm a purely negative duty to avoid believing falsehoods, and claiming that a truth norm only applies in relation to propositions one is consciously considering. I argue that these modified truth norms fail
because they have implausible consequences that make them incapable of explaining justification. There is, however, one modified truth norm, which I will defend, which avoids the problems associated with a prescription to believe all the truths, but does not thereby fail to explain justification. This norm, introduced (but not defended) by Conor McHugh, states that if one has some doxastic attitude about p, then one ought to believe that p if and only if p is true.

Chapter 6 discusses a problem a truth norm faces in explaining suspension of judgement. Some have argued that a truth norm has a problem dealing with suspension of judgement, since suspending judgement about whether p guarantees it that one will fail to have a true belief about whether p. This problem, I claim, leads to a related problem about how a truth norm can explain why one is justified in believing that p only if one has above a certain threshold of evidence in support of p. In order to explain this, I argue, we need to add an extra requirement on belief into our explanation: a requirement that if one fails to conform to the particular truth norm we end up defending, one does so by suspending judgement rather than by believing falsely, i.e. a requirement to err on the side of caution.

Chapter 7 shows how a truth norm can deal with supposedly problematic examples in which having a prima facie unjustified belief is a means to the end of having many other true beliefs, i.e. cases in which a ‘sacrifice for the greater epistemic good’. I argue that the vast majority of these examples are unproblematic because if belief is subject to a truth norm, then justification is proposition-relative, i.e. a belief that p is justified if and only if having that belief is a good means the end of having a true belief that p. Although this deals with the majority of sacrifices for the greater epistemic good, there is still a remaining class of problematic cases. These are self-fulfilling beliefs – beliefs that cause their own truth – such as a belief that people like me, which makes me more sociable, which causes people to like me. In order to explain why self-fulfilling beliefs are unjustified, I argue that we need to further amend the basic truth norm picture by adding a further causal independence condition on justification, i.e. to claim that a belief that p is not justified if it causally dependent on the fact that I believe that p.
Throughout this thesis as a whole, I discuss a large number of norms on belief and other principles. All of the important ones are listed here, as a tool of reference, along with the number of the page in which they are first introduced. I do not defend all of them. I use small caps to indicate the names of norms or principles. I use names rather than numbers to avoid excessive chisholming.¹

1. Truth norms

   OUGHT: One ought to believe that p if and only if p is true. (9)
   CORRECT: A belief that p is correct if and only if p is true. (16)
   GOOD: A belief that p is good if and only if p is true and bad if and only if p is false. (16)
   FITTING: p is a fitting object of belief if and only if p is true. (16)
   ONLY IF: One ought to believe that p only if p is true. (122)
   MAY: One may believe that p if and only if p is true. (122)
   OUGHT NOT: One ought not to believe that p if and only if p is false. (122)
   CONSIDERS: If one consciously considers p, then one ought to believe that p if and only if p is true. (127)
   TRULY BELIEVABLE: If one consciously considers p and p is truly believable, then one ought to believe that p if and only if p is true. (128)
   CONSIDERS-PRESENT: If one is currently consciously considering p, then one ought (at that time) to believe that p if and only if p is true. (130)
   CONSIDERS-PAST: If one has consciously considered p at some time, then one ought (from then on) to believe that p if and only if p is true. (130)

¹ chisholm, v. To make repeated small alterations in a definition or example. “He started with definition (d.8) and kept chisholming away at it until he ended up with (d.8′′′′′′′′′′)” (Dennett and Steglich-Petersen 2008).
DOXASTIC ATTITUDE: If one has some doxastic attitude towards p, one ought to believe that p if and only if p is true. (132)

IF: One ought to believe that p if p is true. (160)

ERR ON THE SIDE OF CAUTION: If one fails to conform to OUGHT with regards to p, one ought not to do so by believing falsely. (171)

ERR ON THE SIDE OF CAUTION-DA: If one fails to conform to DOXASTIC ATTITUDE with regards to p, one ought not to do so by believing falsely. (189)

2. Evidence norms

EVIDENCE: One ought to believe that p if and only if p is supported by one’s evidence. (9)

MORE THAN 50-50: One ought to believe that p if and only if one’s evidence makes p more likely than not-p. (95)

SUFFICIENT EVIDENCE: One ought to believe that p if and only if one has sufficient evidence that p. (95)

EVIDENCE-THRESHOLD: One ought to believe that p if and only if, given one’s evidence, the probability of p is above a particular probability threshold, T. (95)

EVIDENCE-CREDENCE: One ought to proportion one’s credences to the evidence. (96)

PERMISSIVE EVIDENCE: One may believe that p if and only if one has sufficient evidence that p. (124)

SUFFICIENT EVIDENCE-DA: If one has some doxastic attitude about whether p, one ought to believe that p if and only if one has sufficient evidence that p. (143)

3. Other norms on belief

CONSISTENCY: One ought not to have inconsistent beliefs. (9)

DON’T SELF-FALSIFY: One ought not to believe that p if one can reasonably expect the belief that p to be self-falsifying. (51)

KNOWLEDGE: One ought to believe that p if and only if one knows that p. (89)
Avoid Contradiction: If p is contradictory, one ought not to believe that p. (112)

4. Principles

Veritism: A belief is justified if and only if having that belief is a good means to the end of having true beliefs. (10)

Correctness Internalism: If one judges that φ-ing is correct, one is necessarily and inescapably motivated to φ. (44)

Higher Order Belief: If one believes that p, then, absent some kind of irrationality, one believes that one believes that p. (52)

Belief Distribution: If one believes that p and q, then one believes that p and one believes that q. (52)

Truth Connection: A factor F justifies a belief that p if and only if F counts in favour of p being true. (73)

Wide Guidance: A norm N of the form ‘Do X when in C’ can guide S’s behaviour with respect to X only if S following N can make a difference to S X-ing. (107)

Narrow Guidance: A norm N of the form ‘Do X when in C’ can guide S’s behaviour with respect to X only if S following N can make a difference to S X-ing after S has formed a belief as to whether C obtains. (107)

‘Ought’ Implies ‘Can’: If one ought to φ, then it is possible for one to φ. (119)

‘Ought’ Implies ‘Can Satisfy’: If one ought to φ, then it is possible for one to φ while being obligated to φ. (120)

Proposition-Neutral Veritism: A belief is justified if and only if having that belief is a good means to the end of having true beliefs in general. (178)

Proposition-Relative Veritism: A belief that p is justified if and only if having that belief is a good means to the end of believing that p if and only if p is true. (179)

Causal Independence: One has a justified belief that p only if the truth of p is not causally dependent on the fact that one believes that p. (187)
1

INTRODUCTION

It seems to be a truism that the point of having a justified belief is to have a true one. But what does this mean? And is it true? This thesis argues that on a particular understanding of this claim, it is true. In particular, it argues that belief is subject to a truth norm, i.e. a norm or standard that prescribes believing the truth, such as the following norm:

OUGHT: One ought to believe that p if and only if p is true.

Such a truth norm is supposed to provide a way of understanding the claim that the point of having a justified belief is to have a true one. It does so because it is supposed to explain which beliefs count as justified and which do not.

How is a truth norm supposed to explain justification? This becomes clear if one thinks about how someone would go about trying to ensure that they believe in accordance with this norm. One obviously can’t do this directly, i.e. by somehow comparing what one believes with what is true. Therefore, one is going to have to take some indirect means to try and ensure that one’s beliefs are in accordance with this norm. It seems that the means one should take in order to conform to OUGHT would involve following subsidiary epistemic norms that one can follow directly, such as the following:

CONSISTENCY: One ought not to have inconsistent beliefs.

EVIDENCE: One ought to believe that p if and only if p is supported by one’s evidence.

It is then argued that beliefs that are in accordance with these subsidiary norms just are beliefs that we pre-theoretically take to be justified. For this reason, it is claimed that a truth norm is the fundamental norm on belief, i.e. that belief is subject to evidence and consistency norms because belief is subject to a truth norm like OUGHT. This is how a truth norm is supposed to explain epistemic justification. I’ll call this explanation ‘the truth norm account of justification’.

This introductory chapter will give a brief summary of how, in a very abstract sense, the truth norm account of justification is supposed to work. It will not discuss
arguments for such an account, nor will it deal with objections. Those will be dealt with by Part One and Part Two of the thesis respectively.

I’ll proceed as follows. In § 1, I’ll discuss the more general philosophical context within which the truth norm account of justification is situated. Then, in §2, I’ll outline, in a fairly abstract sense, the way in which a truth norm is supposed to explain justification.

1. The truth norm account of justification in context

The truth norm account of justification can be put into context by thinking of it as the conjunction of two commonly made, but separable, epistemological claims: a) the claim that there is an essential relation between justification and truth, and b) the claim that justification is a normative notion. I’ll discuss some background for each of these claims in turn.

1.1. Veritism

It is often claimed that true belief is the goal of epistemic justification, i.e. that the fundamental point of having justified beliefs is to have true ones. Following Alvin Goldman (2001), I’ll call this claim ‘veritism’.

A classic statement of veritism is given by Laurence BonJour:

The basic role of justification is that of a means to truth, a more directly attainable mediating link between our subjective starting point and our objective goal. We cannot, in most cases at least, bring it about directly that our beliefs are true, but we can presumably bring it about directly (though perhaps only in the long run) that they are epistemically justified. And, if our standards of epistemic justification are appropriately chosen, bringing it about that our beliefs are epistemically justified will also tend to bring it about, in the perhaps even longer run … that they are true. (BonJour 1985, 7–8)

From this we can give the following definition of veritism, which builds on BonJour’s language of ‘means’ and ‘ends’:

**VERITISM**: A belief is justified if and only if having that belief is a good means to the end of having true beliefs.

Veritism is not a niche claim; it crosses epistemological fault lines. Many philosophers, who give otherwise very different accounts of justification, claim that there
is *some* essential connection between justification and true belief (for surveys, see Goldman 2001, 53–55; David 2001, 151–52; Berker 2013a, 350–56).

For example, veritist claims have been made by philosophers on both sides of the internalism/externalism debate. One finds veritist claims made by both externalists and internalists about epistemic justification. For example, Goldman, an externalist reliabilist, claims that in order for a belief to be justified it needs to result from a belief-forming process that is *actually* a good means – i.e. a reliable one – to having true beliefs (Goldman 2001). An internalist such as Richard Foley, on the other hand, claims that it is sufficient for a belief to be justified if it’s a good means to having true beliefs *from the believer’s own perspective* (Foley 1993, chap. 1). This shows that both externalists and internalists can be veritists, because they can agree that justification is a means to the end of having true beliefs, but disagree about the sense of ‘means’ involved. Veritism, in this way, has strong bipartisan support.

Furthermore, philosophers have often defended one particular side of an epistemological dispute on the basis that it better accounts for the essential connection between justification and truth. This can be illustrated by reference to the coherentism/foundationalism debate. Arguments for one side in this debate often appeal, either explicitly or implicitly, to the idea that there is an essential connection between justification and truth.

We can see this with regard to a prominent criticism of coherentism about epistemic justification, the so-called ‘alternative systems objection’ to coherentism. This objection claims that coherence is not sufficient for justification because one could have any number of alternative coherent sets of belief about the world. This possibility is illustrated by an example from Russell:

> [T]here is no reason to suppose that only *one* coherent body of beliefs is possible. It may be that, with sufficient imagination, a novelist might invent a past for the world that would perfectly fit on to what we know, and yet be quite different from the real past. (Russell 1912, 71)\(^1\)

Why does this possibility show that coherence does not make for justification? It is because such examples are taken to show that coherence is not *necessarily* a good test of

---

\(^1\) Russell gives this example in order to criticise the coherence theory of truth, but examples like this have also been claimed to undermine coherentism about justification (see Plantinga 1993, 81–82).
truth. Therefore, the alternative systems objection to coherentism implicitly appeals to veritism. Furthermore, an implicit commitment to veritism is evident in the way coherentists typically respond to the alternative systems objection. They argue that a coherent system of beliefs is, in the long run, likely to be truth-conducive (Davidson 1983; BonJour 1985, 169–88; Lehrer 2006).

Furthermore, veritism is also sometimes appealed to by coherentists in order to argue against foundationalism. For example, BonJour has argued against the foundationalist claim that beliefs can be immediately justified — i.e. be justified without being justified by another belief — by appealing to veritism. He makes the veritist claim that “a basic constraint on any account of the standards of justification for empirical knowledge is that there be good reasons for thinking that following those standards is at least likely to lead to truth” (1985, 30). BonJour then argues that this constraint means that it is impossible for a belief to be justified without being justified by another belief, because in order for any belief of mine to meet this constraint, I need another belief to the effect that the first belief has features that make it likely to be true (1985, 32).

However, foundationalists typically have not responded to this kind of argument by denying veritism, but have done so by giving a version of veritism consistent with foundationalism. They typically do this by arguing that BonJour has built into justified beliefs about the world conditions that only really should apply to higher order beliefs that our beliefs about the world are justified. As Pryor puts it, the foundationalist can claim that “[f]or you simply to have a justified belief about the world … that belief may need to have features that make it … likely to be true, but you need not, in addition, be justified in believing it has such features” (Pryor 2001, 102; cf. Alston 1980; Audi 1993).

This is the way in which veritism crosses epistemological fault lines. Veritism has been defended by philosophers who give otherwise very different account of justification, and has been appealed to when deciding between different accounts of justification. While I can’t hope here to outline the ubiquity of veritism within epistemology, this brief summary should suffice to some of the different ways in which it has been understood, and some of the different contexts in which it has been appealed to.

1.2. Justification as a normative notion

As I said above, the truth norm account of justification is a particular version of veritism,
one which is combined with the claim that epistemic justification is a normative concept. I’m now going to outline what this claim means, and give some reasons for thinking that it is correct.

To make the claim slightly weaker, I think it’s more correct to say that there is at least one sense in which justification is a normative concept. There may also be a non-normative sense of justification, but at least one sense of justification is normative and that’s the sense I’ll be concerned with.

What does it mean to say that justification is a normative concept? What I take this to mean is that a belief being justified or unjustified entails facts about what one ought to believe. For example, to call a belief ‘unjustified’ is to say, inter alia, that one ought not to have that belief.

This is supposed to be a distinctively epistemic sense of ‘ought’. Belief may also be subject to moral or pragmatic ‘oughts’, e.g. it may be the case that one ought not to believe that there are genetic determinants of intelligence, because it’s better to treat people as if there weren’t. However, if belief is subject to moral or pragmatic ‘oughts’, these are distinct from the epistemic ‘ought’ that is related to the justification of belief. It is the epistemic ‘ought’ that I’ll be concerned with. The central claim of the truth norm account of justification is that the epistemic ‘ought’ of justification can be explained derivatively, by the fact that belief is subject to a truth norm.

I think this normative sense of justification is commonplace. The term ‘justification’ itself is somewhat of a philosophical term of art, and I don’t mean that it’s commonplace to use the term justification, let alone use it with these normative implications. However, positive and negative assessment of a subject’s beliefs is definitely commonplace, and it’s natural to use ‘justified’ and ‘unjustified’ as catch-all terms for specifying what is common in all these cases of assessment.

The actual terms we use to assess beliefs and beliefs are in fact very varied. We don’t only describe beliefs and believers as ‘justified’ or ‘unjustified’. We also describe them as ‘rational’ or ‘irrational’, ‘reasonable’ or ‘unreasonable’. We also say that someone does or does not have ‘reason’ to think that something is the case. As was the case with justification, use of these terms implies that one epistemically ought or ought not to have the belief in question.

There may be some purpose in some contexts in making explicit distinctions between justification and these other normative epistemic concepts, but I won’t be doing
this. I will be concerned with the most general sense in which beliefs are assessed epistemically, whether positively or negatively, and I'll just be using the terms ‘justified’ and ‘unjustified’ to refer to that.

The variety of our actual epistemic assessment of beliefs is also evident in the fact that we often don’t use general, topic-neutral terms of epistemic assessment like ‘justified’ or ‘rational’. There are also ‘thick’ normative epistemological concepts, which say a bit more. For example, we describe people as ‘careless’, ‘unthinking’, or ‘judgemental’. We call them ‘dupes’, ‘fools’, or ‘idiots'. We describe people – in virtue of the fact that they believe particular things – as ‘dogmatic’, ‘gullible’, ‘deluded’, or ‘biased’. Although these terms have much richer contents than ‘unjustified’ or ‘irrational’, all of them definitely do make a negative normative epistemological assessment of a person in virtue of something they believe.

Furthermore, we don’t even need to use explicitly epistemic vocabulary for this kind of epistemic assessment of a believer to take place. This kind of epistemic assessment seems to be implicit in questions of legal responsibility. This can be illustrated most clearly by outlining how such assessment determines whether one is liable for negligence. If, unbeknownst to me, an act of mine injures you, I can still be legally liable for your injury if I count as negligent. Whether or not I count as negligent seems to turn on normative epistemic assessment of my beliefs. This is illustrated by a foundational case for Scots and English negligence law, *Donoghue v Stephenson*, in which a Scottish woman, Mary Donoghue, became ill after drinking a ginger beer that contained the decomposing remnants of a snail, and sued the manufacturer of the ginger beer, David Stevenson. The case went to the House of Lords, who ruled in favour of Donoghue, because they ruled that it was reasonable for Stevenson to foresee that his failure to take precautions would be likely to cause injury to others. The case therefore set the precedent that a manufacturer’s legal duty of care is not just to those who buys their products, but also to those it would be reasonable to believe would be affected by their products.

This shows that determinations of negligence liability involve normative epistemic assessment. This is made clear in the opinion of Lord Atkin, one of the judges for the case in the Lords:

The liability for negligence, whether you style it such or treat it as in other systems as a species of “culpa,” is no doubt based upon a general public sentiment of moral wrongdoing
for which the offender must pay. But acts or omissions which any moral code would
censure cannot in a practical world be treated so as to give a right to every person injured by
them to demand relief. In this way rules of law arise which limit the range of complainants
and the extent of their remedy. The rule that you are to love your neighbour becomes in
law, you must not injure your neighbour; and the lawyer’s question, Who is my neighbour?
receives a restricted reply. You must take reasonable care to avoid acts or omissions which
you can reasonably foresee would be likely to injure your neighbour. Who, then, in law is my
neighbour? The answer seems to be – persons who are so closely and directly affected by
my act that I ought reasonably to have them in contemplation as being so affected when I
am directing my mind to the acts or omissions which are called in question.2

Lord Atkin’s test for negligence shows that normative epistemic assessment of beliefs is
involved in determining whether someone is negligent or not. According to this test, I
only count as negligent in injuring someone if, in Lord Atkin’s words, I ought to have
them in my contemplation as affected by my actions, i.e. if I ought to believe that my
actions will likely injure them. In this way, epistemic normative assessment is implicit in
determinations of liability for negligence.

Normative epistemic assessment in this way underlies a) the use of general
epistemic terms like ‘justified’ and ‘rational’, b) the use of thick epistemic vocabulary such
as ‘biased’ and ‘deluded’, and c) determinations of legal liability.

I think it’s extremely natural to think the same kind of normative epistemic
assessment is present in all three cases, and that it is captured by speaking in terms of an
epistemic ‘ought’.

It is this epistemic ‘ought’ that a truth norm is supposed to explain, and to
explain in a way that also does justice to the idea that the point of having a justified belief
is to have a true one – i.e. veritism. In this way, the truth norm account of justification is
a combination of veritism and the claim that justification is a normative concept.

Before I move on to outlining the details of how a truth norm is supposed to
explain justification, I’ll end this section by briefly saying why I’m concentrating on
deontic vocabulary like ‘ought’. One could instead attempt to elucidate epistemic
justification – and epistemic concepts more generally – by appealing to alternative
concepts of normative assessment. For example, one could speak in terms of epistemic

2 Donoghue v Stevenson, p.580 [1932] AC 562 (www.scottishlawreports.org.uk/resources/dvs/donoghue-v-
stevenson-report.html).

15
value (see e.g. Kvanvig 2003), or one could speak in terms of epistemic virtues (see e.g. Ernest Sosa 1980; Zagzebski 1996), or vices (Cassam 2016).

Furthermore, some proponents of the truth norm account of justification have preferred truth norms featuring alternative concepts. For example, Ralph Wedgwood has defended the following norm, which features the concept of ‘correctness’:

**CORRECT**: A belief that p is correct if and only if p is true. (Wedgwood 2002, 267)

Conor McHugh has also, at different times, defended the following two norms which are stated in terms of ‘goodness’ and ‘fittingness’:

**GOOD**: A belief that p is good if and only if p is true and bad if and only if p is false. (McHugh 2012, 19–20)

**FITTING**: p is a fitting object of belief if and only if p is true. (McHugh 2014b; McHugh and Way 2016, 584)

The only reason I don’t defend one of these alternative truth norms is because I don’t think we need to. This is because a deontic truth norm can do the job by itself.

Some proponents of the truth norm account of justification, McHugh in particular, have argued that alternative truth norms featuring different normative concepts can get over problems faced by a deontic truth norm.

I’ve argued elsewhere, however, that McHugh’s argument fails to show other truth norms are preferable to a deontic truth norm (Greenberg and Cowie 2016). While I don’t have the space to go into the details of McHugh’s argument here, the problems McHugh raises for a deontic truth norm are the problems I discuss in Chapters 5, 6, and 7. Since I argue in those chapters that a deontic truth norm can get over these problems, for the most part I’ll only discuss truth norms featuring deontic concepts.

2. How a truth norm explains justification

Having put a truth norm in context, I’ll now outline how it is supposed to explain justification. I’ll illustrate this in relation to the truth norm we started with, **OUGH**:

**OUGH**: One ought to believe that p if and only if p is true.

I’m not going to discuss at this stage whether **OUGH** is the formulation of a truth norm we should defend. I’ll discuss this question in Chapter 5, where in the end I’ll defend an
alternative truth norm. It will be simpler, however, to outline how a truth norm is supposed to explain justification by referring to OUGHT.

A truth norm like OUGHT is supposed to explain justification because the means one should take in order to conform to it would involve having beliefs that we pre-theoretically take to be justified. In this way, the truth norm account of justification, like my definition of veritism, appeals to the notions of means and ends.

It might initially seem that this explanation of justification cannot work. Take the proposition that it will rain tomorrow. The truth norm account of justification claims that the means that I should take in order to follow OUGHT will result in me having a justified belief about whether it will rain tomorrow (or being justified in suspending judgement about this).

Unjustified true beliefs and justified false beliefs create immediate problems for this proposal. Unjustified beliefs can be true. For example, I could truly believe that it will rain tomorrow as the result of a lucky guess. If this were the case, I would be believing what I ought to, according to OUGHT. Furthermore, justified beliefs can be false. I may justifiably believe that it will rain tomorrow because of reports of an incoming storm, a storm which – by some freak weather occurrence – is blown south towards France, leaving England untouched. If this were the case, I would not be believing what I ought to, according to OUGHT. How then is a truth norm like OUGHT supposed to explain justification?

In order to answer these questions we need to focus on two different aspects of the way a truth norm is supposed to explain justification. First, we need to focus on the kind of norm OUGHT is, and the position we are in with respect to that norm. In particular, we need to recognise that a truth norm is an objective norm, i.e. a norm for which one can’t tell directly whether one is in accordance it. Because of this, the truth norm account of justification claims that we should follow subsidiary norms as a means of conforming to a truth norm. I’ll discuss this aspect of the truth norm account of justification in §2.1.

Secondly, a truth norm account of justification needs to specify that it is not just any means of trying to conform to OUGHT that is supposed to result in having justified beliefs. Rather, it is only the good means of conforming to OUGHT or the means one should take to conform to OUGHT. The truth norm account of justification has to involve this claim to give a plausible explanation of justification. However, it requires spelling out
what the ‘good’ means of conforming to a truth norm are. I’ll consider some different ways of doing this in §2.2.

Before I do this, however, I first want to note two things. Firstly, I just want to flag that I’ll predominantly be discussing how a truth norm can explain justification. While proponents of a truth norm do claim that it can explain justification (see e.g. Engel 2004, 88–90; Wedgwood 2008, 4–6; McHugh 2011, 371; Whiting 2013a, 185), they also sometimes claim that a truth norm explains other positive epistemic statuses. Sometimes they claim that a truth norm can explain what it is for a belief to be epistemically rational (Wedgwood 2002; 2013a; Boghossian 2005, 211). Sometimes they claim it can explain what epistemic reasons for belief are (Shah 2003, 471; 2006; Engel 2013a, sec. III; McHugh 2012, 9; 2014b).

There may be, as I’ve already said, purpose in some contexts in making explicit distinctions between these concepts. I don’t think, however, there is reason to do this when thinking about the explanatory role of a truth norm. Therefore, while I’ll be talking about justification, I think everything I say could be said about epistemic rationality or reasons for belief. I won’t, however, discuss whether a truth norm can explain knowledge in addition to justification. Defenders of a truth norm sometimes also claim that it can explain knowledge (see e.g. Wedgwood 2002, secs. 5–6). I’m going to set this question aside, and concentrate purely on justification. The main reasons why are a) reasons of space, and b) because I don’t think we should assume that there is any simple relation between knowledge and justification. The post-Gettier literature should at least make us question that assumption.

Secondly, I also want to flag that I won’t discuss an alternative way in which a truth norm is claimed to explain justification. The most popular version of the truth norm account of justification is the one I’m outlining, i.e. one which appeals to subsidiary norms one should follow as a means to the end of conforming to a truth norm (see e.g. Wedgwood 2002, 276; Boghossian 2003, 38–39; 2005, 211–12; Shah 2003, 471; Shah and Velleman 2005, 520; Engel 2004, 82; 2013a, 207; Millar 2004, 43–47; 2009, 145–46). However, some other philosophers respond to the questions raised by justified false beliefs by claiming that, despite appearances, there are no justified false beliefs, i.e. all justified beliefs are true (see e.g. Littlejohn 2012, chaps. 3–4; Steglich-Petersen 2013b). I won’t be discussing such a view because I’m not willing to drop the assumption that there are justified false beliefs. Claiming that are no justified false beliefs would, I believe,
result in a picture that would completely fail to track our intuitive understanding of when someone counts as justified, and would fail to track the kinds of normative epistemic assessment I outlined above. Furthermore, as I'll outline in Chapter 2, I think the only real support we can give for a truth norm is to show that it can explain our pre-theoretical intuitions about justification. Therefore, I’m going retain the assumption that there are justified false beliefs, and pursue a version of the truth norm account of justification that can accommodate that assumption.

2.1. Objective and subjective norms on belief

OUGHT is not the kind of norm I can follow directly. This is because OUGHT is an objective norm. Following Paul Boghossian (2003, 38–39; 2005, 211–12), we can think of an objective norm as a norm for which there is no direct way of telling whether one is in accordance with it. Some different examples of objective norms Boghossian gives are the following two norms about how one ought to play the stock market:

BUY LOW: One ought to buy stocks when their price is low

SELL HIGH: One ought to sell stocks when their price is high

When buying and selling stocks, there is no direct way of telling whether one is doing so in accordance with these norms, i.e. one can’t tell directly whether the stock one is buying will continue to drop in price, or whether it’s about to stabilize and start rising.

Therefore, in order to try and play the stock market in accordance with these two objective norms, I need to follow some other subjective norm. In this context, a subjective norm is a) a norm for which there is a direct way of telling whether I’m in accordance with it, and b) a norm which I follow only as a means to the end of trying to be in accordance with one or other objective norms.

One subjective norm I could follow in order to conform with BUY LOW and SELL HIGH is the following:

FINANCIAL TIMES: One ought to buy and sell what the Financial Times recommends to buy and sell.

FINANCIAL TIMES is a norm for which there is a direct way of telling whether I am in accordance with it, i.e. I can check the newspaper. Secondly, following that norm is not an end in itself, but only a means of attempting to trade in accordance with the objective
norms BUY LOW and SELL HIGH.

Actual traders will probably follow different subjective norms. As Boghossian puts it, “some will use rules based on technical indicators, others will use rules based on fundamentals” (2005, 211). The fact remains that following such norms is not an end in itself, but a more directly accessible means of attempting to trade in accordance with BUY LOW and SELL HIGH; and by saying that the stock traders’ norms are ‘a more directly accessible means’ is not to say that following them is easy – just that following them directly does not require omniscience.

OUGHT is also an objective norm. Given that I’m not omniscient, I can’t tell directly whether I’m believing in accordance with this norm. Consider again the proposition that it will rain tomorrow. I can’t just check whether my attitude about whether it will rain tomorrow is true, in the way that I can check whether my stock purchases match the recommendations in the Financial Times. Therefore, as with BUY LOW and SELL HIGH, in order to try and ensure that my attitude about whether it will rain tomorrow is in accordance with OUGHT, I will have to take some means to that end.

A clear statement of this appeal to means is given by Ralph Wedgwood. He does this by appealing to the idea of ‘aiming at truth’. He suggests that if one is aiming to conform with a truth norm like OUGHT, the means one would take is to adopt certain policies or rules for belief formation and revision:

Suppose that you are considering a proposition \( p \), and are literally ‘aiming’ to conform to this fundamental epistemic norm. That is, you are aiming to believe \( p \) if and only if \( p \) is true. Clearly, you will not end up achieving this aim simply because it is your aim. You will have to do something in order to achieve this aim. That is, you will have to do something by means of which (if all goes well) you will achieve this aim. Presumably, in order to achieve this aim, you must revise your beliefs in certain ways when you are in certain conditions, and revise your beliefs in other ways when in other conditions. (Wedgwood 2002, 276)

What rules should I follow, according to Wedgwood, if I am aiming to believe in accordance with OUGHT? First, I should follow rules for forming and revising beliefs that won’t lead me to believing falsely. As Wedgwood puts it, I should follow rules that are ‘reliably error-avoiding in the circumstances’ (Wedgwood 2002, 277). However, in order to conform to OUGHT, I shouldn’t follow just any rules that are reliably error-avoiding. For example, I shouldn’t follow a rule that prescribes me to suspend judgement about everything. This is because such a rule, while reliably error-avoiding, also ensures that I’ll
never believe truly about any matter, and thus never have any beliefs that succeed in being in accordance with OUGHT.

Therefore, in order to conform with OUGHT, I should follow rules that are not so reliably error avoiding that they prevent me from having true beliefs. As Wedgwood puts it; I should follow rules that not only are reliably error avoiding in the circumstances, but rules that are also ‘no more restrictive than necessary’ (Wedgwood 2002, 278).

These rules state, in a very general and abstract sense, the means one should take to conform to OUGHT. Proponents of the truth norm account, including Wedgwood, also suggest more specific means one should take to this end. In fact, proponents of the truth norm account claim that the means one should take include following the two subsidiary norms I mentioned at the beginning of this chapter:

**CONSISTENCY:** One ought not to have inconsistent beliefs.

**EVIDENCE:** One ought to believe that p if and only if p is supported by one’s evidence.

Ensuring that one’s beliefs are consistent is one means Wedgwood suggests one should take to the end of conforming to OUGHT (Wedgwood 2013b, 124–25). This is because if one’s beliefs are inconsistent, they cannot all be true.

Basing one’s beliefs on evidence – i.e. believing that p if and only if one p is supported by one’s evidence – is also suggested as one of the means one should use to this end (Boghossian 2003, 39; 2005, 211; Engel 2013a, 207). This is because, proponents of the truth norm account claim, evidence in favour of a hypothesis just is something that counts in favour of that hypothesis being true (Engel 2013a, 207). Therefore, in order to conform to OUGHT, one needs to base one’s beliefs on evidence.

Both Wedgwood’s rules for belief formation and revision and the norms CONSISTENCY and EVIDENCE serve as subjective norms, in the sense I outlined above. First, they are rules and norms for which, unlike OUGHT, one can tell directly whether one is in accordance with them. Again, that’s not to say that this would be easy, but just that one doesn’t have to be omniscient to do it. Secondly, one doesn’t, according to the truth norm account, follow these rules and norms as an end in itself, but only as a means of trying to ensure that one believes in accordance with OUGHT.

These won’t be the only subjective norms involved in the truth norm account of justification. Nor do I think we’ve said enough about the particular subjective epistemic
norms the truth norm account will have to appeal to in order to explain justification. In order to fully defend the truth norm account of justification, much more will have to be said on these matters, and I’ll go into much more detail – especially with regards to evidence norms – in Chapter 6.

In summary, the central claim of the truth norm account of justification is that these subsidiary subjective norms are means to the end of conforming with an objective truth norm. It is in virtue of this means-end relation that justification is explained by a truth norm like OUGHT. This is because a subject who follows these subjective norms has beliefs which we pre-theoretically take to be justified; but those subjective norms were just the means the subject should take to conform with OUGHT. Therefore, through the means-end relation, the fact that belief is subject to OUGHT explains what it is for a belief to be justified. Justification is, as it were, the means to truth.

2.2. What are the ‘good means’ of conforming to a truth norm?

The definition of veritism I started with did not just claim that any means one might take to the end of believing truly results in having justified beliefs; rather, it claimed that only good means of believing truly result in justified beliefs:

VERITISM: A belief is justified if and only if having that belief is a good means to the end of having true beliefs.

‘Good’ needs to be in the definition because an unjustified belief can look like it is in some sense a means to having true beliefs. For example, say that I believe that I will get bad news this week because I read as much in my horoscope and in the end I do happen to get bad news this week. Having the belief that I will get bad news this week may be, in some sense, a means to the end of having true beliefs. However, the belief is not justified – according to veritism – because having that belief does not look like a good means to having true beliefs.

If the truth norm account of justification is going to work, it’s going to need the same qualification. This is because, in the example above, believing I will get bad news this week because I read it in my horoscope does look to be, in some sense, a means to the end of conforming to OUGHT. But it doesn’t look like a good means of conforming to OUGHT. One only has justified beliefs, according to the truth norm account, if one follows subsidiary norms that are good means of conforming to OUGHT.
In this way, the truth norm account of justification needs to appeal to a notion of ‘good means’ or the means one ‘should’ take to conform to a truth norm. It’s therefore an explanatory burden on the proponent of the truth norm account to unpack what ‘good’ means are. Since throughout the thesis I’ll be discussing what are good means of conforming to truth norms and what are not, I’ll quickly discuss here a few different ways of unpacking ‘good means’, and in the end suggest what I take to be the correct way.

It will be illuminating to start by discussing two ways of understanding ‘good means’ which clearly won’t work. First, one could understand ‘good means of conforming to OUGHT’ as ‘means which will in fact lead to one conforming with OUGHT’. On such an account, one would count as justified if one followed a subsidiary norm that in fact leads to conformity with OUGHT. This is clearly far too strong, because it has the result that all true beliefs are justified and no false beliefs are justified. Therefore, if ‘good means’ of conforming to OUGHT are understood like this, OUGHT will fail to explain justification.

Secondly, one could understand ‘good means of conforming to OUGHT’ as ‘means which one believes will lead to one conforming to OUGHT’. On such an account, one would count as justified if one followed a subsidiary norm that one believed would lead to conformity with OUGHT. This is clearly far too weak, because merely believing that following an epistemic norm will lead to one believing truly does not look sufficient for being justified in following that norm. Presumably, some people do believe that believing what their horoscope says will lead to them believing truly, but that doesn’t by itself make them justified in believing what their horoscope says.

These two ways of understanding ‘good means’ provide two unsatisfactory extremes; the former being too strong, the latter too weak. The correct unpacking of what are ‘good means’ will have to lie in between these two extremes.

A third way of understanding what ‘good means’ is given by Wedgwood. As we saw earlier, Wedgwood thinks the means one should take to conform to OUGHT are to follow rules for belief formation and revision that are reliably error-avoiding but no more restrictive than necessary. However, that was a slight oversimplification. Wedgwood doesn’t claim that one should follow rules that are in fact reliably error-avoiding in order to conform to OUGHT; instead, he claims that one should follow those rules which “it is rational for you to believe to be reliably error-avoiding” (Wedgwood 2002, 278–79, my
italics). Therefore, according to Wedgwood’s account the ‘good means of conforming to OUGHT’ are ‘the means that it would be rational to believe will lead to conformity with OUGHT’.

I don’t think, however, that Wedgwood’s specification of ‘good means’ will work either. This is not because it is either too strong or too weak; rather, it is because it is problematically circular. Wedgwood takes himself, as we’ve outlined, to be giving an explanation of what it’s rational to believe in terms of a truth norm (he elsewhere claims that we can give exactly the same explanation of justification (2008, 4–6), so it’s clear that he views ‘rational’ and ‘justified’ fairly interchangeably). If we have to appeal to the rules it is rational to believe are reliable in order to show how a truth norm can successfully explain rationality, then this explanation is problematically circular.

Wedgwood does note this circularity, but claims that it is unproblematic:

Roughly, rational beliefs are beliefs that either result from, or (in the case of background beliefs) amount to, one’s following a rule or set of rules that it is rational for one to believe to be reliable. Admittedly, this specification of this property is not fully non-circular, since it uses the term ‘rational’. But for our purposes we do not need a fully non-circular specification. This specification is enough to show that the universal norm of rational belief is explained by the fundamental norm of correct belief. (Wedgwood 2002, 282)

It’s not clear, however, that Wedgwood is right to think that this circularity is unproblematic. If Wedgwood is correct that we have to appeal to a notion like rationality in order to show how a truth norm can succeed in explaining justification, then surely the correct conclusion is that the concept of rationality – rather than a truth norm – is doing the explaining.

A fourth possible way of understanding ‘good means’ is to give an externalist reliabilist reading of it. Such an account would understand ‘good means of conforming to OUGHT’ as ‘means which would in fact reliably lead one conforming to OUGHT’. This may seem to strike the right middle ground, since it a) does not entails that all and only true beliefs are justified, and b) merely believing the horoscope to be reliable is not sufficient, on this account, for believing the horoscope to be a good means of conforming to OUGHT.

However, I don’t think this understanding of ‘good means’ will work either. This is because it is unstable and will collapse back into the first, overly strong understanding of ‘good means’, i.e. ‘means which will in fact lead to one conforming with OUGHT’. For,
Once one accepts the externalism implicit in this understanding of ‘good’ means, one cannot then deny that following a norm which will in fact lead to conformity with ought is a better means of conforming to ought than following a norm which merely reliably leads to conformity with ought. To illustrate, consider the following two norms:

**True Horoscope:** One ought to believe what the horoscope says when and only when the horoscope speaks the truth.

**Ignore the Horoscope:** One ought to never believe what the horoscope says.

If one is giving a purely externalist account of what are ‘good means’ of conforming to ought, then it’s hard to deny that following True Horoscope is a better means of that than following Ignore the Horoscope. Of course, following True Horoscope is not something that’s really available to us to do, since we’re not in a position to discern when the horoscope is going to get it right by luck. But that shouldn’t matter if we’re giving an externalist understanding of ‘good means’. Therefore, an externalist reliabilist reading of ‘good means’ collapses into an understanding of ‘good means’ that is overly strong.

Before I move on, I should quickly note that this does not mean that reliabilism about justification is necessarily wrong, just that someone with reliabilist sympathies shouldn’t appeal to a truth norm.

A fifth way of understanding ‘good means’ – and the one I’ll prefer – takes its cue from the problem we found in the reliabilist reading. We should understand ‘good means of conforming to ought’, I suggest, as ‘those means which a subject is in a position to discern that they will lead to conformity to ought’. In other words, a norm on belief only counts as good means of conforming to ought if a subject is in a position to tell that it is.

Understanding ‘good means’ in this way gets over the problems with the other accounts. Firstly, it looks initially like it won’t be too strong or too weak – at least not in the way that the first two ways of understanding good means were found to be.

Secondly, it is not problematically circular in the way Wedgwood’s account is. I do admit that the question of whether or not one is a position to discern something is an epistemic question, and one which we might not be able to a reductive answer to. However, my understanding of good means does not appeal to the very concept we are using a truth norm like ought to explain, i.e. justification or rationality, which we saw to be the problem with Wedgwood’s account.
Thirdly, it will not, as the reliabilist account of ‘good’ means did, collapse into the overly strong account of good means. This is because we’re not always in a position to discern whether believing something will lead to us conforming to OUGHT. This means, for example, that following TRUE HOROSCOPE does not count as ‘good means’ on this understanding because, as I said, we’re not in a position to discern when the horoscope gets it right by luck.

This last point does mean that – in opting for my understanding of ‘good means’ – the resulting truth norm account will have to involve a modicum of epistemic internalism, because it results in justification being relative to what one is in a position to discern. However, this is an extremely moderate amount of internalism, and doesn’t entail anything about necessarily having access to one’s justification, grounds, or reasons for belief. Therefore, I don’t think it commits the truth norm account of justification to any heavyweight version of epistemic internalism.

In this brief canvassing of the different ways to understand what are ‘good means’ of conforming to a truth norm, I’ve concluded that a subjective epistemic norm only counts as a good means of conforming to OUGHT if a subject is in a position to tell that it is. In the broader context of how a truth norm explains justification, this means one counts as justified if one follows epistemic norms which one is in a position to discern that they will lead to one conforming to OUGHT.

Therefore, when we think about whether or not different subjective epistemic norms – such as consistency and evidence norms – are good means of conforming to a truth norm, we need to think about it in this fashion; i.e. whether or not they will lead to conformity with a truth norm in a way that is discernible to a subject. When I discuss whether or not particular evidence norms are good means of conforming to particular truth norms – as I will do throughout this thesis – it should be understood in this way.

3. Conclusion

This concludes my summary of the way a truth norm is supposed to explain justification, i.e. by generating subsidiary norms one should follow as a means to the end of conforming to a fundamental truth norm.

Before we move on, I want to note two things. Firstly, in this summary I purely viewed a truth norm as part of first-order epistemology. As we’ll see in the next chapter, many philosophers think a truth norm plays an explanatory role in the philosophy of
mind as well. I'll go into some of the psychological phenomena a truth norm is taken to explain in the next chapter, though I'll be sceptical about whether such phenomena are best explained by a truth norm. Secondly, I haven't taken myself to have provided any arguments in defence of a truth norm in this chapter. I have merely sketched the explanatory role it is sometimes taken to play within epistemology. However, sketching this explanatory role provides the necessary backdrop to the argument I'll give in defence of the truth norm. This is because, as I'll also argue in the next chapter, the best argument for the claim that belief is subject to a truth norm is that such a norm genuinely can succeed in explaining epistemic justification.
Part One of this thesis lays the groundwork for an epistemological argument in defence of a truth norm. It does so by distinguishing between two broad ways of arguing for a truth norm. The first way – which I’m sceptical of – is to argue for a truth norm on the basis of its explanatory role in the philosophy of mind. The second way – which I favour – is to argue for a truth norm on the basis of its explanatory role in epistemology.

Chapter 2 argues against the first of these two ways of arguing for a truth norm, by discussing and rejecting a number of arguments that claim a truth norm is required to explain various different psychological phenomena. Furthermore, these psychological arguments for a truth norm are unnecessary; if a truth norm can play an explanatory role within epistemology, that by itself would provide sufficient reason to think that belief is indeed subject to a truth norm.

Chapter 3 outlines one central explanatory role a truth norm can play within epistemology. It argues that a truth norm can explain what unifies justification, i.e. it can explain what the various different methods and considerations that justify belief all have in common.

I do not at this stage deal with any problems a truth norm might face. All Part One concludes is that a truth norm does not have any explanatory power in the philosophy of mind, but could have explanatory power in epistemology. In this way, all Part One attempts to show is that an epistemological argument for a truth norm has promise, and that a psychological argument does not. Part Two of this thesis follows through on that promise, by arguing that there is a defensible explanation of justification in terms of a truth norm that can deal with the variety of problems that have been raised for such an explanation.
This chapter discusses two broad ways of arguing for the claim that belief is subject to a truth norm. The first way – which I’ll express scepticism about – is to argue that belief is subject a truth norm on the basis of the explanatory role of such a norm in the philosophy of mind. Such arguments typically claim that we need to think of the attitude of belief as constitutively subject to a truth norm in order to explain some psychological feature of belief. Although arguments of this kind have been a very popular way of arguing for a truth norm, I am sceptical about whether they can work. I’ll argue that either these psychological features of belief can be explained by much weaker claims than a truth norm, or that a truth norm can only explain these features given very strong theoretical assumptions.

The second broad way of arguing for a truth norm – which I’ll favour – is on the basis of its explanatory role in epistemology. We should not try to give independent reason for thinking belief is subject to a truth norm, a norm which we then go on to use to explain justification. We should instead think of belief as subject to a truth norm because that can do justice to our pre-theoretical understanding of justification, much as one might argue for a moral norm or theory because it does justice to our pre-theoretical understanding of what we ought to do. The best argument for a truth norm is not one that appeals to the philosophy of mind, but one within epistemology.

I’ll proceed as follows. I’ll first go through a series of psychological features of belief that philosophers have claimed should make us conclude that belief is subject to a truth norm. I’ll discuss claims that a truth norm explains, in turn, what distinguishes belief from other attitudes (§1), the normativity of content (§2), the transparency of doxastic deliberation (§3), the impossibility of believing at will (§4), and Moore’s paradox (§5). I’ll argue that none of these features give good reason to think that belief is subject to a truth norm. Many of these arguments defend a truth norm stated in terms of ‘correctness’. Because of this, I’ll speak in terms of correctness often in this chapter, rather than in the deontic terms that I prefer.

I’ll then put forward an alternative way of arguing for a truth norm, on the basis of its explanatory role in epistemology (§6). This chapter itself will not provide the full
argument for a truth norm on the basis of its explanatory role in epistemology – that will be provided by the thesis as a whole. All I’m arguing here is that the prospects for arguing for a truth norm in this fashion are better than the prospects for arguing that a truth norm is constitutive of the attitude of belief. Therefore, as well as providing a critical survey of the arguments that have been given in defence of a truth norm, this chapter also sets the scene for my epistemological argument – given in the following chapters – that belief is subject to a truth norm.

1. Can a truth norm explain what distinguishes belief from other attitudes?

The first argument for a truth norm I’ll discuss purports to show that we need to appeal to the claim that belief is subject to a truth norm in order to explain what distinguishes a belief that p from other attitudes one could have towards p. Other features of belief, this argument claims, are insufficient to explain what distinguishes belief from other attitudes. A clear expression of this thought comes from Michael Lynch:

[T]he fact that truth is a norm of correctness for believing is part of what distinguishes believing from other cognitive attitudes. Imagining, assuming, and hoping, for example, are each governed by norms – assumptions can be justified or not, imaginings can be sharp or vague, hopes can be rational or irrational. But neither imagining that p, assuming that p, nor hoping that p are properly evaluated in terms of truth. Believing is. (Lynch 2009, 229)

Why should one think that we have to appeal to a truth norm in order to distinguish belief from other attitudes?

A prominent argument for this claim comes from David Velleman. Velleman argues that we need to appeal to a truth norm to distinguish belief from other attitudes because no other feature of belief suffices to do so. Velleman does this by arguing against the commonly made claim that belief’s motivational or functional role distinguishes belief from other attitudes, i.e. the claim that what distinguishes belief from other attitudes is that believing that p involves the disposition to act in a way that would satisfy one’s desires if p were true (see e.g. Ramsey 1926, 69; Braithwaite 1932, 30; Stalnaker 1984, 15). According to this ‘motivational’ view of belief, what distinguishes believing that p from imagining that p, assuming that p, hoping that p, and all other attitudes one could have towards p is that believing that p involves – whereas other attitudes do not – being disposed to act in a way that would satisfy one’s desires if p were true.

Velleman argues that the motivational view of belief is false. He does so by
arguing that other attitudes towards p involve the disposition to act in a way that would satisfy one’s desires if p were true. In particular, Velleman claims that imagining that p involves this disposition. Velleman argues for this claim by appealing to the example of make-believe, in which, he claims, imagining that p disposes a child to act in a way that would satisfy his desires if p were true. Velleman claims that “when [a child] imagines that he is an elephant, he is disposed to behave as if he were an elephant” (2000, 256), e.g. by stomping around, making elephant like noises, etc. This means, Velleman claims, that imagining that p involves the same disposition that the motivational view claims is distinctive of belief (2000, 256). Velleman concludes that we need something else to distinguish believing that p from imagining that p. And he claims that what in fact distinguishes these two attitudes is what Lynch above suggests: believing that p is – whereas imagining that p is not – an attitude that is correct if and only if it’s true (Velleman 2000, 277; Shah and Velleman 2005, 497–98).

Has Velleman shown that belief is subject to a truth norm? I am sceptical. Firstly, not just any disposition to act as if p were true is the one which the motivational view claims is distinctive of belief. Specifically, it’s only the disposition to act in a way that would satisfy one’s desires if p were true. This specific disposition is required because it helps distinguish those who really believe that p from, for example, actors or liars who make it seem that they believe that p. Those who make it seem that they believe, say, that God exists because they are lying or acting in a film are disposed to act as if it’s true that God exists, but they are not disposed to act in a way that would satisfy their desires if it were true that God exists. So it is unclear why we can’t say the same about the child who imagines that he is an elephant as part of an episode of make-believe, i.e. that he is disposed to act

---

1 This a slight oversimplification of Velleman’s view. Velleman initially (in his 2000) concluded that what distinguished belief from imagining is that a belief is an attitude one adopts towards p with the aim of doing so if and only if p is true (where that aim is either a conscious intention or the subconscious goal-directed activity of one’s cognitive system). Later, and in response to criticism of his initial view (in Shah 2003), Velleman concludes that belief is not distinguished by the psychological feature of being adopted with the aim of doing so if and only if p is true; rather, Velleman now holds that belief is distinguished by the normative feature that it is correct if and only if it’s true (Shah and Velleman 2005). However, it’s clear that Velleman still regards himself as having refuted the motivational view of belief with the arguments I’ve outlined (Shah and Velleman 2005, 497–98). Therefore, since I am only concerned with the claim that belief is distinguished by being subject to a truth norm, I’m going to cut to the chase, and treat Velleman as if he argues for a truth norm directly.
as if he were elephant, but that he is not disposed to act in a way that would satisfy his desires if it were true that he were an elephant.

The fact that someone who merely imagines that they are an elephant is not disposed to act in a way that would satisfy their desires if it were true that they are an elephant can be seen by thinking about what it would be like to actually act in a way that would satisfy one’s desires if it were true that one were an elephant. This point is made clearly by Lucy O’Brien:

If I somehow could come to believe – rather than have a delusion – that I (LOB) am an elephant then, it would seem to have a very different motivational role to the corresponding imaginings. I would probably be trying to find ways to resign from my job, break it to my family, buy a new bed and so on. I feel motivated to do none of these things when I imagine I am an elephant. (O’Brien 2005, 59)

The same point is made about an imagining with a different content by Kathrin Glüer and Asa Wikforss:

If I imagine that I am an eagle, and I desire to fly, I am not thereby motivated to jump of a cliff. (Glüer and Wikforss 2013b, 145)

This is what it would be like to really have to have the disposition to act in a way that would satisfy one’s desires if it were true that one were an elephant or an eagle, but this is not the case with someone who merely imagines that they are an elephant or an eagle. In this way, just as with actors and liars, someone who merely imagines that they are elephant does not have the disposition to act which the motivational view claims is distinctive of belief.

Velleman does recognise what we’ve shown – that imagining and believing do not motivate action in this same way. But he claims this is merely the result of the fact that imagining that p is usually accompanied by a belief that not-p, and therefore accompanied by the stronger countervailing disposition to act in a way that would satisfy one’s desires if not-p were true (Velleman 2000, 272). Therefore, the fact that the child also believes that he is not an elephant means that he has a stronger countervailing

---

2 O’Brien actually makes this point in the process of a criticism of Velleman that gives more credence to Velleman’s claims about the motivational force of imagination than I do, but the point she makes can be adapted to my criticism as well.
disposition to act in a way that would satisfy his desires if it were true that he was an elephant. That is why, according to Velleman, he is not disposed to do *everything* he would do if he really were an elephant.

This response looks unconvincing. If Velleman were right, then imagining that I’m an elephant should at least give me *some* motivation to do the things O’Brien lists (resign from my job, buy a bigger bed, etc.), motivation which is then outweighed by my motivation not to do these things given by my belief that I’m not an elephant. But imagining that I’m an elephant doesn’t seem to give me *any* motivation to do these things. Therefore, it doesn’t look like Velleman is correct that imagining shares the same disposition to act that belief does.

Therefore, since the motivational view of what distinguishes belief from other attitudes is not undermined in the way that he claims it to be, Velleman has not shown that we *need* another account of what distinguishes belief from other attitudes. Therefore, the claim that belief is subject to a truth norm does not find any support from this source.

Furthermore, even if we grant there is some other argument against the motivational view of belief, and we do need to, as Velleman claims, appeal to the claim beliefs are correct if and only if true, that still looks insufficient to defend the *kind of* truth norm that is needed for the truth norm account of justification.

This is because it is unclear that a truth norm required to distinguish belief from other attitudes would have sufficient normative force, as I’ll put it, to explain justification. What I mean by a norm having ‘normative force’ is that the norm entails that one *ought* to perform an action or have some attitude.

Some uses of ‘correct’ clearly do *not* have normative force in this sense. To adapt an example of Gideon Rosen’s, there are correct and incorrect ways to play Mozart’s C-Major sonata. However, the correctness of certain ways of playing the sonata does not *by itself* entail anything about how one ought to play it in a particular way. Rosen illustrates this with the following example (he makes the point in relation to whether one has *reason* to play Mozart’s C-Major sonata correctly, but we can make an analogous point in relation to whether one *ought* to play the sonata correctly):

Your daughter is practicing the piano, but she’s getting frustrated. So you decide to cheer her up with a laughable performance of her piece. You start in on Mozart’s C-major sonata, but you play a C# where a B should be as the first note of the second measure. Now you
knew all along that to play the C# you would be to play the piece incorrectly. Did this judgment provide you with a reason not to play the note? To the contrary, it seems plausible that your judgment of correctness was practically neutral in itself. In the normal case when you aim to play the piece correctly, the thought that C# would be the wrong note provides you with a reason not to play it. But when you have every reason (as in this case) to play the piece incorrectly, the very same thought has the opposite valence. (Rosen 2001, 620–21)

Something being a ‘correct’ playing of Mozart’s C-major sonata, by itself, has no normative force, i.e. it can’t in any explain by itself why one ought to play it that way. I’ll call this the deflationary sense of ‘correct’.

Recall the role a truth norm is supposed to play in explaining justification. In order for a truth norm to explain justification it seems that a truth norm which belief is subject to must have normative force, i.e. it needs to entail, in particular, facts about what one ought to believe. If true belief is correct only in this deflationary sense, that doesn’t look like it can, by itself, explain justification, because that doesn’t entail any facts about what one ought to believe.

In the context of the arguments we’re discussing in this section, it seems that the opponents of a truth norm can claim that true belief being correct in the deflationary sense is sufficient to explain what distinguishes belief from other attitudes (again, assuming that there is good reason to reject the motivational view of belief, and that we need to appeal to belief being correct if and only if true in other to distinguish belief from other attitudes).

Of course, there may be something the proponent of a truth norm can say in response here, something about why in order to distinguish belief from other attitudes we need to appeal to a sense of correctness for belief that, unlike the correctness of Mozart’s C-Major sonata, does have normative force. But that is an additional argument that the proponent of a truth norm who takes this line would have to make, and is bound to involve additional theoretical assumptions, which are liable to make an argument for a truth norm along these lines less and less plausible.

In conclusion, we should be sceptical of an argument that claims a truth norm is needed to distinguish belief from other attitudes. Firstly, the argument that the alternative motivational account fails to distinguish belief from other attitudes is implausible. Secondly, even if we would need to appeal to belief being correct if and only if true in order to distinguish belief from other attitudes, supposing that belief is correct in a merely deflationary sense would suffice to distinguish belief from other attitudes.
Another reason sometimes given for thinking that belief is subject to a truth norm is that it can explain why meaning or mental content is normative. The claim that meaning is normative is most famously made by Saul Kripke, in his discussion of Wittgenstein’s rule-following considerations. Kripke claims that what constitutes my meaning something by a word or sign – for example, meaning addition by ‘+’ – cannot just be how I have been disposed to use that word or sign. This, Kripke argues, leaves out the normative element:

Suppose I do mean addition by ‘+’. What is the relation of this supposition to the question how I will respond to the problem ‘68+57’? The dispositionalist gives a descriptive account of this relation: if ‘+’ meant addition, then I will answer ‘125’. But this is not the proper account of the relation, which is normative, not descriptive. The point is not that, if I meant addition by ‘+’, I will answer ‘125’, but that, if I intend to accord with my past meaning of ‘+’ I should answer ‘125’. Computational error, finiteness of my capacity, and other disturbing factors may lead me not to be disposed to respond as I should, but if so, I have not acted in accordance with my intentions. The relation of meaning and intention to future action is normative, not descriptive. (Kripke 1982, 37)

Many have found it plausible that meaning is normative in the way Kripke outlines here, and have also claimed that mental content is normative in the same kind of way (see e.g. Brandom 1994; Gibbard 2003; Boghossian 2003). It has been suggested by some that the normativity of meaning or mental content can be explained by the fact that belief is an attitude that is constitutively subject to a truth norm (Velleman 2000, 244–45; Boghossian 2003; 2005).

The most sustained attempt to do this is made by Paul Boghossian. Boghossian argues that the normativity of mental content can be explained derivatively by the fact that belief is essentially subject to a truth norm. He bases his explanation on a claim that our understanding of content is derivative of our understanding of the content of belief. Boghossian outlines this argumentative strategy in the following passage:

If our grasp of the notion of content were somehow to depend in a privileged and asymmetric way on our grasp of the concept of belief, then our only route to the notion of a contentful state would be through our grasp of a constitutively normative notion, and – although we would have arrived at this result in a way not envisioned by its proponents, still – that would be enough to substantiate the claim that content itself is normative, in spirit if
Boghossian proceeds to argue that our understanding of mental content in general is in fact dependent on our understanding of the concept of belief. He does this by arguing that our understanding of contentful attitudes other than belief – desires, hopes, fears – is dependent on our understanding of belief. He concludes this first on the basis of concluding that the concept of desire is dependent on the concept of belief, and then draws a general moral from this conclusion.

How does Boghossian argue that the concept of desire is dependent on the concept of belief? This claim, Boghossian claims, is plausible because it is possible for a person to have the concept of a belief without having the concept of desire, but not *vice versa*. Boghossian says that “it does seem possible for someone to have the idea of accepting a content as true without having any idea of what it would be to desire a content to be true” (2005, 214). On the other hand, according to Boghossian, it doesn’t look possible for someone to have the concept of a desire without having the concept of belief:

> Could someone understand the idea of wanting the world to be a certain way, but have no idea at all of what it would be to take it to be a certain way, to accept its being a certain way? This does seem bizarre. Don’t I have to think of someone as having some beliefs about how things are, in order to coherently think of them as having wants about how things should be? (Boghossian 2005, 215)

Because of this, Boghossian concludes that the concept of desire is dependent on the concept of belief. Furthermore, he thinks this kind of argument will likely show that the concepts of other attitudes – being glad, sad, angry, conflicted, etc. – are dependent on the concept of belief in the same way as the concept of desire is (2005, 216). We therefore should conclude, according to Boghossian, that our understanding of any attitude having a content is dependent on our understanding of a belief having a content.

What’s the upshot of this? Does this provide an argument for a truth norm? It’s not clear whether Boghossian takes his explanation of the normativity of content to provide an argument for a truth norm, but he certainly claims that belief is essentially subject to a truth norm when giving that explanation. In any case, I’ll consider whether Boghossian’s explanation does in fact support the claim that belief is subject to a truth norm.
Should we think that belief is subject to a truth norm because it explains how and why mental content is normative? There are a number of reasons to be sceptical of such an argument.

Firstly, one may doubt Boghossian’s claim that the concept of desire is dependent on the concept of belief, both on conceptual and empirical grounds. It has been disputed on conceptual grounds by Alexander Miller. Miller claims that the conceptual relationship between belief and action should make us conclude that there is a relation of \textit{interdependence} between the concepts of belief and desire, rather than concluding, as Boghossian does, that there is an asymmetric dependence of the concept of desire on the concept of belief:

\begin{quote}
Grasping the concept of belief requires grasping that beliefs potentially issue in action, which in turn involves grasping that beliefs can lead to action by combining with desires. Would we say that someone with no inkling of this nevertheless possessed the concept of belief? I think not. (Miller 2008, 236–37)
\end{quote}

Boghossian’s claim of dependence has also been disputed by Glüer and Wikforss, on empirical grounds, who claim that there is empirical evidence that the concept of desire is conceptually \textit{independent} from the concept of belief:

\begin{quote}
There is evidence that the thesis, in fact, is false – today it is widely accepted in developmental psychology that children use the concept of desire in explanatory contexts long before they acquire the concept of belief. (Glüer and Wikforss 2009, 40)
\end{quote}

I’m not here going to defend either of these two alternative ways of understanding the relationship between the concepts of belief and desire. However, highlighting these two alternatives shows how strong Boghossian’s claim is. Boghossian does not provide any arguments against these alternatives, and he would have to do so in order to defend his claim of dependence. Such an argument in defence of Boghossian’s claim of dependence would be bound to involve further theoretical assumptions which would themselves need defence. This is liable to make an argument for a truth norm along these lines less and less plausible.

Secondly, even if we grant Boghossian his claim of dependence, it doesn’t seem that belief needs to be essentially subject to a \textit{truth} norm in particular in order to explain the normativity of content. Belief being subject to \textit{any} norm could play the explanatory role Boghossian outlines. For example, if a truth norm could play this explanatory role,
so could a norm that requires one’s beliefs ought to be *rational*, or a norm that requires one’s beliefs to be based on *evidence*, or a norm that states that one ought to only believe that p if one *knows* that p. Some of these alternative norms are less controversial than a truth norm. Therefore, if we need to appeal to a norm on belief in order to explain the normativity of content — a claim we should be sceptical of in any case — it would seem to be dialectical preferable to appeal to an alternative norm on belief to play this explanatory role, rather than appealing to a truth norm. (As we’ll later see, Boghossian’s argument shares this flaw with a number of other arguments for a truth norm, which also cannot support a truth norm over other, less controversial norms on belief).

Thirdly, the claim that mental content is normative is a hugely controversial thesis, which itself has faced much criticism (see e.g. Papineau 1999; Dretske 2001; Hattiangadi 2007; Glüer and Wikforss 2009). It is dialectically unwise for a proponent of a truth norm to yoke their defence of a truth norm account to such a controversial thesis.

In summary, if there’s an argument for a truth norm given by Boghossian, it’s an argument that a truth norm explains a hugely controversial thesis by way of a contestable premise. And it wouldn’t show that belief is subject to specifically a *truth* norm in any case. The prospects for such an argument do not look good.

### 3. Can a truth norm explain the transparency of doxastic deliberation?

Another reason given for thinking that belief is subject to a truth norm is that it can explain a phenomenon that has been called ‘transparency’. Velleman and Nishi Shah have argued that belief is subject to a truth norm on this basis (Shah 2003; Shah and Velleman 2005).

What Shah and Velleman call ‘transparency’ is a feature of first-personal deliberation about what to believe, which they describe as follows:

> The deliberative question *whether to believe that* p *inevitably gives way to the factual question* *whether* p, because the answer to the latter question will determine the answer to the former.

> That is, the only way to answer the question *whether to believe that* p *is to answer the question* *whether* p. (Shah and Velleman 2005, 499)

Why does deliberating about whether to believe that p inevitably result in thinking about whether p? Shah and Velleman claim that this fact is best explained by belief being constitutively subject to a truth norm. They argue for this by, for the most part, arguing that another explanation cannot work. In particular, they reject a possible explanation of
transparency in terms of dispositional facts about belief. One might think that transparency could be explained by belief being a state one is disposed to be in if one has evidence that it is true. If believing that p involves such a disposition, that might seem to explain transparency, i.e. to explain why deliberating about whether to believe that p inevitably results in thinking about whether p. On this account, deliberation about whether to believe that p inevitably results in thinking about whether p because a belief that p just is an attitude one is disposed to have if one has evidence that p is true.

Shah and Velleman argue that this explanation cannot work because a disposition of this kind could only explain transparency, they claim, if it were implausibly strong. The disposition needs to be so strong because the transparency of belief is a feature which doesn’t just show that evidence as to whether p is true is one factor that is relevant to answering the question of whether to believe that p; rather, it shows that evidence as to whether p is true is the only relevant factor. Therefore, in order to explain transparency, the disposition to be responsive to evidence of truth would have to exclude any influences on belief formation other than evidence, i.e. a belief that p would have to be an attitude that one is disposed to have if and only if one has evidence that p is true (Shah 2003, 462–64; Shah and Velleman 2005, 500–501).

However, it is implausible, Shah and Velleman claim, that belief involves such a strong disposition. This is because we have to accept, Shah and Velleman claim, that beliefs are sometimes influenced by factors other than evidence of truth. For example, beliefs are sometimes be influenced by pragmatic factors, such as in cases of wishful thinking (Shah 2003, 461–62; Shah and Velleman 2005, 500). Consider the case of a parent who believes that their child’s bad marks are down to the teacher’s animus, not because of evidence, but because of their desire for their child to be clever and successful. Such a belief is influenced by factors other than evidence of truth, as are many others.

Because of this, Shah and Velleman conclude that belief cannot involve such a strong disposition to be responsive to evidence of truth, i.e. to be an attitude that one is disposed to have if and only if one has evidence that p is true. This is because, if believing were to essentially involve this strong disposition, that would implausibly rule out beliefs like that of the child’s parent, which are influenced by pragmatic factors, from counting as beliefs (Shah 2003, 461–62; Shah and Velleman 2005, 500).

However, it’s only this strong disposition to be responsive to evidence of truth,
Shah and Velleman claim, that can explain transparency. This is because, as we’ve already said, when deliberating about whether to believe that p, evidence as to whether p is true is the *only* relevant factor. A weaker disposition to be responsive to evidence of truth – such as the disposition to be responsive, *among other things*, to evidence of truth – would allow beliefs produced by wishful thinking to count as beliefs. However, this weaker disposition cannot explain transparency, because evidence of truth is the *only* factor relevant to answering the question of whether to believe that p (Shah 2003, 462; Shah and Velleman 2005, 501).

Because of this, Shah and Velleman conclude that dispositional facts about belief are incapable of explaining transparency, and claim that we should instead opt for a *normative* explanation. In particular, in order to explain transparency, according to Shah and Velleman, we need to think of belief as an attitude that is constitutively subject to a truth norm. Shah and Velleman outline this explanation in the following passage:

> When one deliberates whether to have an attitude conceived as a belief that p, one deliberates about an attitude to which one already applies the standard of being correct if and only if p is true, and so one is already committed to consider it with an eye exclusively to whether p. (Shah and Velleman 2005, 501)

As Shah and Velleman outline in this passage, this is supposed to explain transparency because when one is deliberating about whether to believe that p, they claim, one applies the standards that are implicit in belief, and those standards make evidence of truth what is *exclusively* relevant.

However, it’s also an explanation which, according to Shah and Velleman, does not rule out the products of wishful thinking from counting as beliefs. This is because, according to their explanation, it’s only in the context of first-personal deliberation about what to believe that the subject applies the truth norm, and thus commits themselves ‘to consider it with an eye exclusively to whether p’. This, Shah and Velleman claims, allows pragmatic considerations to influence belief formation in other contexts (Shah and Velleman 2005, 501).

In summary, Shah and Velleman claim that belief being constitutively subject to a truth norm is the best explanation of the transparency of belief; it’s claimed to be an explanation which both explains why evidence of truth is the only relevant factor in doxastic deliberation, but does not thereby rule out the products of wishful thinking from counting as beliefs.
Is this a good argument for the claim that belief is subject to a truth norm? As with the previous arguments, there are reasons to be sceptical.

A first problem is that Shah and Velleman are arguing against a strawman. The claim that a truth norm is the best explanation of transparency is argued for, by Shah and Velleman, by claiming that transparency cannot be explained by one dispositional account of belief. Furthermore, it’s a very simple dispositional account, according to which the sole determinant of whether or not an attitude is one of belief is whether or not it’s responsive to evidence.

This looks like a strawman. There are more sophisticated dispositional accounts of belief which Shah and Velleman could discuss, some of which may not fail to explain the transparency of belief. Firstly, there is the dispositional account we discussed above, according to which belief is distinguished by a disposition to lead to actions. Secondly, there also seem to be a wide variety of possible dispositional accounts of belief which we could adapt from the literature on causal reductions of intentionality. This literature aims to explain intentionality in terms of causal relations in between thinkers and the objects their thoughts are about (for a summary, see Loewer 1999). A central problem of this literature is to explain how one can have false thoughts, because it is not as straightforward how the content of such thoughts can be explained in terms of causal relations in between thinkers and the things their thoughts are about. In response to this problem, various philosophers have given more sophisticated attempts at causal reduction which claim that different subsets of the causal relations are the content-determining ones (see e.g. Stampe 1977; Stalnaker 1984, chap. 1; Dretske 1986; Fodor 1987, chap. 4). It seems that there is the avenue of appealing to parallels of the moves made in this literature in order to develop further more sophisticated dispositional accounts of belief.

Given the many options for more sophisticated dispositional accounts of belief, Shah and Velleman are arguing against a strawman. Some more sophisticated dispositional account may be able to both a) explain the transparency of belief, and b) not rule out the products of wishful thinking from counting as beliefs. At the very least,

3 Shah and Velleman reject such an account, but they do so for Velleman’s reasons we discussed and rejected above (Shah and Velleman 2005, 497–98), so this is not a good reason to ignore whether such an account could explain transparency.
given the fact that Shah and Velleman have only discussed a single, very simple dispositional account, their claim that no dispositional account of belief can do these two things is unwarranted.

A second problem with Shah and Velleman’s purported explanation of transparency is that it seems to rely on an implausibly strong form of internalism about motivation. This should make us doubt that a norm in particular could do the explanatory work that Shah and Velleman think a truth norm is doing.

This is a point made by Asbjørn Steglich-Petersen. The phenomenon of transparency which Shah and Velleman are trying to explain is, we said above, the inevitable transition from the question of whether to believe that p to the question of whether p is true. Given that, Steglich-Petersen points out that, on Shah and Velleman’s explanation, “if transparency is produced by the norm of belief, this norm motivates one necessarily and inescapably to act in accordance with it” (Steglich-Petersen 2006, 507).

The claim that there is this kind of necessary relationship between a norm and motivation is a form of internalism about motivation. In particular it’s a version of motivational judgement internalism about correctness, which we can formulate as follows:

**CORRECTNESS INTERNALISM:** If one judges that φ-ing is correct, one is necessarily and inescapably motivated to φ.

The problem for Shah and Velleman is that this is an extremely strong version of motivational judgement internalism. It is not plausible that there is such a strong link between norms and motivations. Steglich-Petersen makes this point with regards to practical norms. For example, it is probably true that if one promises to φ, then one ought (at least pro tanto) to φ. But it seems that one could make a promise, judge that promising is subject to this norm, but not be inescapably motivated to keep one’s promise (Steglich-Petersen 2006, 506–7). Such a strong link between norms and motivation is implausible, and therefore, it’s implausible, Steglich-Petersen claims, to think of transparency as being explained by a norm:

[T]he motivation stemming from the thought that true beliefs are correct has to be so strong, if it is to do the desired explanatory work, that it is implausible to regard it as motivation stemming from acceptance of a norm at all. (Steglich-Petersen 2006, 506)
Furthermore, Steglich-Petersen’s claim can be further justified, I suggest, by considering the forms of motivational judgement internalism actually defended by moral philosophers. Those who claim that there is some necessary connection between judging an action to be morally right (or good, correct, or what one ought to do) and being motivated to perform that action typically claim that this connection is, in various ways defeasible. Some claim that the moral judgement necessarily results in some motivation to act, but not necessarily overriding motivation; others claim that moral judgement necessarily results in motivation to act in ideal circumstances, or if one is rational (for a survey, see Svavarsdottir 1999, sec. 1). A representative example of this kind of motivational judgement internalism is put forward by Michael Smith:

[Agents who judge it right to act in various ways are so motivated, and necessarily so, absent the distorting influences of weakness of the will and other similar forms of practical unreason on their motivation. (Smith 1994, 61)]

These qualified versions of motivational judgement internalism are typically defended, as the passage from Smith indicates, because a version of internalism without such qualifications would have the implausible consequence that it would be impossible to judge that something is right but fail to do so out of weakness of the will, deep depression, etc.

However, the kind of internalism about motivation that Shah and Velleman’s explanation of transparency would have to rely on — i.e. CORRECTNESS INTERNALISM — is a version of internalism that claims that there is an unqualified necessary connection between judgement and motivation, because it claims that judging something to be correct inescapably motivates one to do it. In this way, the kind of internalism Shah and Velleman would have to rely on is the exact kind of connection between normative judgement and motivation that motivational internalists do not defend. This gives support to Steglich-Petersen’s claim that the motivational internalism they would have to rely on is implausibly strong.

A response to Steglich-Petersen’s criticism is given by Pascal Engel. Engel argues that a truth norm does not, as Steglich-Petersen claims, need to motivate necessarily and inescapably, and thus Shah and Velleman’s explanation does not need to be committed to an implausibly strong form of internalism. This is because it is possible, Engel claims, to recognise that a truth norm prescribes believing that p, but to fail to believe that p. This happens, Engel claims, in cases of irrationality:
Just as cases of akrasia or acedia can arise where the agent considers the norm but does not follow it, cases where the norm of truth is considered by the agent but is not followed can arise. (Engel 2007, 199)

Engel claims that such cases show that Steglich-Petersen is mistaken in thinking that a truth norm must, on Shah and Velleman’s account, motivate necessarily and inescapably.

I don’t think, however, that Engel’s response works. Engel is right that if the kind of cases of epistemic akrasia he envisages exists, then a truth norm does not motivate necessarily and inescapably. However, this does not vindicate Shah and Velleman’s argument for a truth norm.

This is because there is a dialectical problem with Engel’s response, pointed out by Andrei Buleandra (2009, 328–31). If these kinds of cases of epistemic akrasia exist, that also casts doubt on the claim that belief exhibits the phenomenon of transparency; in particular, it casts doubt on Shah and Velleman’s claim that there is an inescapable and inevitable transition from the question of whether to believe that p to whether p is true. As Buleandra points out, if one can judge that believing that p is correct (because p is true) but not thereby come to believe that p – as Engel claims is possible – that means that transition is not inevitable (Buleandra 2009, 329). But it was precisely the inevitability of this transition that meant, as we saw earlier, that the alternative dispositional explanation of transparency failed, and we needed, according to Shah and Velleman, to appeal to a truth norm. Therefore, if that cases of epistemic akrasia exist, then the phenomenon to be explained – transparency – does not exist. Therefore, if Engel’s right, then Shah and Velleman’s argument for a truth norm fails, because there is no phenomenon that needs to be explained by a truth norm.

A third and final problem with Shah and Velleman’s argument is that even if the above problems can be dealt with, it is unclear that a truth norm in particular is needed to explain transparency. Again, it seems that some alternative epistemic norms could provide an equally good explanation. Consider the following evidential norm:

**EVIDENCE**: One ought to believe that p if and only if p is supported by one’s evidence.

Proponents of a truth norm typically do accept that belief is subject to evidential norms like this one, but also claim that it is subject to evidential norms because it is subject to the fundamental truth norm (see Shah 2006; cf. Boghossian 2003, 39; Whiting 2010, 221–22;
Engel 2013a, 207; McHugh 2014b). But there is an alternative view, defended most prominently by Richard Feldman and Earl Conee, according to which an evidential norm is the sole fundamental norm on belief, and not a derivative norm that is explained by the more fundamental truth norm (Conee and Feldman 1985; Feldman 1988; 2000). Even ignoring the previous two problems I’ve raised for it, Shah and Velleman’s argument it would not, I suggest, support the claim that a truth norm – rather than an evidential norm – is the fundamental norm on belief.

The key reason why this is the case is because Shah and Velleman describe the purported phenomenon of transparency in a way that stacks the deck in favour of it supporting a truth norm. Shah and Velleman claim that the phenomenon to be explained – transparency – is the inevitable transition between the question of whether to believe that p to the question of whether p is true. There seems room to doubt whether this ‘truth-centric’ way of describing the phenomenon is correct. In particular, a more evidence-centric way of describing it could be correct, i.e. that the inevitable transition is actually between the question of whether to believe that p to the question of whether there is evidence that p. If this latter evidence-centric way of describing the phenomenon exhibited by belief is correct, then it would be an evidential norm, rather than a truth norm, that would better explain the phenomenon. Furthermore, it seems to me to be an open question as to whether the phenomenon of transparency is best described in the truth-centric or evidence-centric way.

This undermines Shah and Velleman’s argument for a truth norm. For Shah and Velleman to have shown that a truth norm is required to explain transparency, it needs to be the case the truth-centric way of describing the phenomenon is uniquely correct. But it’s not clear that this is the case, and Shah and Velleman have given no reason for thinking it is. Therefore, even if Shah and Velleman can find some way of getting over the other problems I’ve raised for their argument, it wouldn’t support the claim that belief is subject to a truth norm, rather than another norm.

4. Can a truth norm explain the impossibility of believing at will?

The next argument for a truth norm that I’ll discuss is drawn from Bernard Williams’ discussion of the impossibility of believing at will. Williams himself doesn’t appeal to the impossibility of believing at will in order to argue for a truth norm; rather, he argues that the impossibility of believing at will should be explained by the fact that belief ‘aims at
truth’. No philosopher has, to my knowledge, argued explicitly for a truth norm on the basis that it’s impossible to believe at will. However, a truth norm has been claimed to be, by a number of its defenders, as a way to ‘interpret’ Williams’ claim that belief aims at truth (see e.g. Wedgwood 2002, 267; Engel 2013b, 32). Therefore, I’m going to discuss whether the impossibility of believing at will can provide support for a truth norm, i.e. whether the impossibility of believing at will is a psychological feature that is best explained by belief being an attitude which is constitutively subject to a truth norm.

Williams claims in the following passage that the fact that belief aims at truth explains why it’s not just a contingent, but a necessary truth that one cannot form a belief at will:

[II]t is not a contingent fact that I cannot bring it about, just like that, that I believe something, as it is a contingent fact that I cannot bring it about, just like that, that I’m blushing. Why is this? One reason is connected with the characteristic of beliefs that they aim at truth. 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. (B. A. O. Williams 1970, 148)

I want to discuss whether there is a good argument for a truth norm that is parallel to the one Williams gives in this passage. In other words, does the fact that it is impossible to ‘just bring it about’ that I believe that p mean that belief is constitutively subject to a truth norm. Such an argument would claim that if I were able to acquire an attitude without any regard to its truth, I could not seriously think of such an attitude as a belief, because I would be self-consciously violating the norm, implicit in belief, that prescribes believing that p if and only if p is true.

Again, I am sceptical about whether an argument like this can work. A first reason for scepticism is that a norm on belief doesn’t really look like the appropriate kind of thing to explain the impossibility of believing at will. I’ll outline two reasons why.

Firstly, if a norm on belief is doing explanatory work here, then it’s going to have to be the case that a self-conscious violation of that norm is impossible, given explanandum is the impossibility of believing at will. Given this, an argument that a truth norm explains the impossibility of believing at will would, just as Shah and Velleman’s argument, rely on an implausibly strong form of motivational judgement internalism.

Secondly, a norm looks inappropriate in explaining the possibility of believing at
will because a norm that states that I ought to $\phi$ seems to presuppose that I am, at least to some degree, free to $\phi$ or not to $\phi$. If one has no freedom about whether or not one $\phi$-ies – i.e. if one will inevitably either $\phi$ or fail to $\phi$ – then it seems inappropriate to say that one ought to $\phi$. This is a thought expressed by Rousseau’s remark that “[i]f force compels obedience, there is no need to invoke a duty to obey” (1762, Bk. I, Ch. 3).\(^4\)

The fact that norms presuppose some degree of freedom makes a truth norm look like the wrong kind of thing to explain the impossibility of believing at will. We would be explaining the fact that believing is not up to us, by appealing to something else – a norm on belief – that presupposes that believing is, in some sense, up to us.

The proponent of a truth norm could claim at this point that we do have the requisite freedom about what we believe. Many different accounts have been given of the way in which we count as free or responsible with respect to our beliefs (and how that differs from the kind of freedom and responsibility we have with respect to actions) (see e.g. Steup 2000; Heller 2000; Ryan 2003; Hieronymi 2006; McHugh 2014a). But if the proponent of a truth norm claims that we do have the requisite freedom about what we believe, that undermines the claim that there genuinely is a phenomenon – the impossibility of believing at will – that needs to be explained by thinking of belief as constitutively subject to a truth norm. Therefore, for these two reasons we should be sceptical of the claim that a truth norm would be the right kind of thing to explain the impossibility of believing at will.

However, even if these issues can be dealt with, an argument for a truth norm along these lines still faces the issue which many psychological arguments for a truth norm face – the fact that other norms on belief would do an equally good job of explaining the phenomenon. Even if the impossibility of believing at will is, despite appearances, appropriately explained by belief being constitutively subject to a norm, it wouldn’t need to be specifically a truth norm that would be needed to do the explanatory work. If a truth norm could explain the impossibility of believing at will, then a wide variety of norms on belief – e.g. a knowledge norm, an evidential norm, or a rationality norm – could equally well explain the impossibility of believing at will. This is because the only work a truth norm was doing in the above explanation was being a norm on

\(^4\) For an in-depth discussion of these issues, see Railton 1999, esp. 322-25.
belief which, because it’s constitutive of belief, one couldn’t self-consciously violate. This was supposed to explain why it’s impossible to believe, say, something one wants to believe. All the other norms I’ve mentioned could – if they were constitutive of belief – play exactly the same role as a truth norm in this explanation. Therefore, there are no prospects for an argument for in particular a truth norm on the basis that it’s impossible to believe at will.

5. Can a truth norm explain Moore's paradox?

Moore’s paradox is another feature of belief that might be thought to support a truth norm. Consider the following propositions:

- It’s raining, but I don’t believe that it is.
- I believe it’s raining, but it’s not.

It would be absurd to ever believe or assert these propositions. However, it’s possible for either of these propositions to be true; they are not self-contradictory. Given this, it seems that there needs to be some explanation of why it is absurd to believing or assert a Moore-paradoxical proposition.

One might think that in order to explain this absurdity we need to appeal to the fact that belief is subject to a truth norm. Both Williams and Peter Railton suggest that the absurdity of asserting a sentence that expresses a Moore-paradoxical proposition reflects the fact that belief aims at truth (B. A. O. Williams 1970, 137; Railton 1994, 72–73). As I’ve said above, defenders of a truth norm often claim that in claiming that belief is subject to a truth norm, they are a giving a normative interpretation of the fact that belief aims at truth. So one might think that there are prospects for another argument for a truth norm here, one that argues that we need to think of belief as constitutively subject to a truth norm in order to explain the particular absurdity of asserting (or believing) a Moore-paradoxical proposition.

An argument of this kind is sketched by John Gibbons (2013, 2–3). Although he does not end up defending a truth norm, Gibbons regards Moore’s paradox as providing at least prima facie reason for thinking that belief is subject to a truth norm. Therefore, we can use his argument as an exemplar to assess whether or not Moore’s paradox provides support for a truth norm.

Moore’s paradox provides support for a truth norm because, according to
Gibbons, it shows that the requirement encoded in a truth norm is non-optional in a way that other requirements – such as the requirements of etiquette – are not:

The truth norm on belief is not similarly optional. Suppose that you try to reject the norm, know the relevant condition obtains, and know you’re not doing what it tells you to do. These are the things to think:

(O) It’s raining, but I don’t believe it.
(C) I believe it’s raining, but it’s not.

The first is an admission that you’ve violated the rule that tells you to believe the truth, and the second is an admission that you’ve violated the rule that tells you to avoid believing what’s false. Of course, (O) is the omissive form of Moore’s Paradox, and (C) is the commissive form. Whatever else you want to say about Moore’s Paradox, it’s fairly clear that the attitudes expressed by (O) and (C) are incoherent, irrational, or maybe even absurd. So we cannot simply and self-consciously reject the truth norm on belief, at least not in the sense that we can simply and self-consciously reject the requirements of etiquette. (Gibbons 2013, 3)

Does the absurdity of believing or asserting a Moore-paradoxical proposition support, in the way that Gibbons indicates, the claim that belief is subject to a truth norm? Again I think there is reason to be sceptical.

Let’s grant that Moorean absurdity should be explained as an absurdity arising from the absurdity of self-consciously rejecting the requirements of a norm on belief. (This seems to be a claim that we could deny; it is denied by both those who claim that Moorean absurdity is fundamentally an absurdity associated with speech and those who claim than Moorean absurdity is the manifestation of practical irrationality (see Green and Williams 2007, chap. 1 for a catalogue of the different approaches to Moore’s paradox)). Even if we accept this, it’s still unclear that we need to appeal to specifically a truth norm in order to explain Moorean absurdity. This is because it’s again the case that alternative norms on belief can explain the absurdity of believing a Moore-paradoxical proposition. Furthermore, some of these alternative norms that can do the requisite explanatory work in this case are much less controversial than a truth norm.

Firstly, consider the following norm put forward by John Williams (2012, 1132):

**DON’T SELF-FALSIFY:** One ought not to believe that p if one can reasonably expect the belief that p to be self-falsifying.

This norm could equally well explain the absurdity of believing a Moore-paradoxical
proposition. This is because beliefs in Moore-paradoxical propositions are, by their very content, self-falsifying, i.e. they are proposition that are necessarily false if believed. This means that beliefs in Moore-paradoxical propositions are beliefs one can reasonably expect to be self-falsifying, and thus prohibited by DON’T SELF-FALSIFY. According to this explanation, DON’T SELF-FALSIFY – and not a truth norm, as Gibbons suggests – is the norm which one can’t self-consciously reject.

This is problematic because DON’T SELF-FALSIFY looks much weaker and less controversial than a truth norm. Many philosophers who would not want to accept that one ought to believe something if and only if it’s true would nevertheless accept that one ought to never have a self-falsifying belief. Therefore, since the absurdity of believing a Moore-paradoxical proposition can be explained by a much weaker norm than a truth norm, it does not provide support for a truth norm in the way Gibbons indicates.

To hammer the point home, we can make much the same point with another alternative norm on belief. Consider the following norm:

CONSISTENCY: One ought not to have inconsistent beliefs.

Again this norm on belief is much weaker and less controversial than a truth norm.

But CONSISTENCY can also explain the absurdity of believing a Moore-paradoxical proposition if we make two fairly weak assumptions. The first is an assumption about higher order belief.

HIGHER ORDER BELIEF: If one believes that p, then, absent some kind of irrationality, one believes that one believes that p.

HIGHER ORDER BELIEF claims that if one doesn’t believe that one believes that p when one believes that p, then one is necessarily irrational.

The second assumption is that belief distributes over conjunction:

BELIEF DISTRIBUTION: If one believes that p and q, then one believes that p and one believes that q.

Given these two assumptions, CONSISTENCY is sufficient to explain the irrationality involved in believing a Moore-paradoxical proposition. Assume I believe the following Moore-paradoxical proposition:
It’s raining, but I don’t believe that it’s raining.

Given higher order belief, this means that unless I’m irrational, I believe the following proposition:

I believe that (it’s raining, but I don’t believe that it’s raining)

But given belief distribution, having both this higher order belief and the original first-order Moorean-belief entails that I simply having straightforwardly inconsistent beliefs. Given belief distribution and the fact that I have the first-order Moorean belief, I believe the following proposition:

I don’t believe that it’s raining

Given belief distribution and the fact that I have the higher order Moorean belief, I believe the following proposition:

I believe that it’s raining.

These beliefs are straightforwardly inconsistent, and thus prohibited by consistency.

Therefore, given the two fairly weak assumptions higher order evidence and belief distribution, the alternative norm on belief consistency can explain why I’m inevitably irrational if I believe a Moore-paradoxical proposition. Either I don’t have the higher order belief that I believe a Moore-paradoxical proposition and I’m thereby irrational according to higher order belief, or I do have the higher order belief and I’m thereby irrational according to consistency.

This, again, casts doubt on Gibbons’ claim that the absurdity of believing a Moore-paradoxical proposition provides support for a truth norm. This is because consistency is much weaker and less controversial than a truth norm. Many philosophers who would not want to accept that one ought to believe something if and only if it’s true would nevertheless accept that one ought not to have inconsistent beliefs. And although I’ve had to appeal to some other assumptions to use consistency to explain Moorean absurdity, these assumptions were themselves very weak. Therefore, this explanation that appeals to consistency provides further support for the claim that a truth norm is not the weakest norm on belief we need to appeal to in order to explain the absurdity of believing a Moore-paradoxical proposition.
In giving these alternative explanations, I don’t mean to say that DON’T SELF-FALSIFY and CONSISTENCY are alternatives to a truth norm as candidates for being the fundamental norm on belief that explains justification. They are not. Neither DON’T SELF-FALSIFY nor CONSISTENCY could explain justification on their own in the way that proponents of a truth norm suggest that a truth norm can.

This does not, however, make a difference to the point I’ve been making. My point is that a truth norm is not the weakest possible norm on belief that could explain the absurdity of believing a Moore-paradoxical proposition. This is true completely independently of whether or not DON’T SELF-FALSIFY and CONSISTENCY can explain justification by themselves.

We can make essentially the same point if it is claimed that belief is subject to DON’T SELF-FALSIFY and AVOID CONTRADICTION because it is subject to a truth norm. This may be true, but it does not make a difference to my point here. The claim that, within the truth norm account of justification, DON’T SELF-FALSIFY and AVOID CONTRADICTION can both be explained by a truth norm makes no difference to my claim that a truth norm is not the weakest norm on belief that can explain the absurdity of believing a Moore-paradoxical proposition. If my claim is true, then Moore's paradox does not support the claim that belief is subject to a truth norm in particular.

6. A better way: belief is subject to a truth norm because that explains epistemic justification

I’ll end this chapter by suggesting a different kind of approach to arguing for a truth norm. We should not, I suggest, first argue for a truth norm because of its supposed explanatory role in the philosophy of mind, and only then use it to explain justification. Rather, we should argue for a truth norm on the basis of its explanatory role in epistemology. The best argument for a truth norm, in this way, comes from within epistemology.

This distinction between these two alternative ways of arguing for a truth norm has not been clearly made in the literature. This is probably down to the fact that many proponents of a truth norm both think that it is constitutive of belief, and that it can explain justification. This has meant that the idea of a truth norm is typically run together with the idea that belief is constitutively subject to such a norm. I know of only one proponent of a truth norm – Steglich-Petersen – who clearly distinguishes between the two: “[O]ne can coherently reject the essential normativity of mental states, including
belief, while accepting that beliefs can be guided by a truth norm” (Steglich-Petersen 2013a, 279, fn. 2).

Steglich-Petersen is correct that these two ideas are distinct. One can – and, according to me, should – think belief is subject to a truth norm without thinking that it is constitutively so. I therefore want to offer an alternative way of arguing for a truth norm that only appeals to its explanatory role within epistemology.

I don’t want to claim that this way of arguing for a truth norm is entirely novel. Some philosophers do at times argue for (Wedgwood 2002) or against (Gibbons 2013, Ch. 4) a truth norm on purely epistemological grounds. More generally, discussion of a truth norm often includes some discussion of the explanatory role of a truth norm within epistemology (see e.g. Boghossian 2003, 38–39; 2005, 211–12; Engel 2004, 82; 2013a, 207; Whiting 2010, 220–22). However, this is typically regarded as only one part of the story about a truth norm. I want to suggest that it can and should be regarded as the whole story.

In this section, I’ll outline and briefly defend this alternative methodology. A truth norm can be defended, I suggest, by showing that it provides a unifying explanation and encapsulation of our intuitive understanding of justification. I’m not concerned in this chapter with showing that a truth norm in fact can provide this kind of explanation of justification – this thesis as a whole aims to provide support for that claim. My concern here is just to show that if a truth norm can provide this kind of unifying explanation of justification, that by itself would be sufficient to support the claim that belief is subject to a truth norm.

If I’m right about this, that makes the arguments for a truth norm on the basis of its explanatory role in the philosophy of mind dialectically pointless. We don’t need to provide independent support for a truth norm which we then go on to use to explain justification; that it can explain justification is by itself reason enough to think that belief is subject to a truth norm. Given the shortcomings I’ve outlined in the arguments for a truth norm on the basis of its explanatory role in the philosophy of mind, it’s better if we can argue for a truth norm without having to appeal to them. If I’m right, we don’t have to.

As I’ve said, I intend to defend a truth norm by showing that it provides a unifying explanation and encapsulation of our intuitive understanding of justification. However, our intuitive understanding of justification includes some elements that are
more concrete and others that are more abstract. I am going to suggest that a truth norm can provide a unifying explanation of all of these elements of justification.

The more concrete aspect of our intuitive understanding of justification is an intuitive understanding of when someone counts as justified in believing something in a particular case and when someone does not. A truth norm explains this because, I'll suggest, it prescribes believing that p in cases in which it looks intuitively plausible that one would be justified in believing that p. It also, I'll suggest, prescribes suspending judgement in cases in which it looks intuitively plausible that one would be justified in doing so. A truth norm does this, recall, because one has to follow subsidiary norms as a means to the end of conforming to it. This is the way in which a truth norm explains how the more concrete elements of our intuitive understanding of justification cohere together. I'll argue for this claim, in Ch. 4–7, mainly in response to a variety of criticisms of a truth norm approach which, in effect, claim that a truth norm should be rejected because it gives the wrong prescriptions in particular cases.

In addition, our intuitive understanding of justification includes some more abstract, semi-theoretical elements. Firstly, I think the claim I've called ‘veritism’ – the claim that there is some essential connection between justification and truth – is a more theoretical element of our intuitive picture of justification. Since a truth norm is, as I suggested, one way to understand veritism, a truth norm can do justice to this element of our intuitive understanding of justification.

A second more theoretical element is the idea that epistemic justification is unified. This is the idea that all the manifold different things that can justify a belief are - at some level of description - all of the same type. This is manifest in the fact that one can weigh seemingly diverse considerations for and against a belief against each other. I'll go into this in more detail in Chapter 3, where I'll argue that a truth norm can unify epistemic justification in this way.

In summary, I am going to argue that we should think belief is subject to a truth norm because it can give a unifying explanation of how these different elements of our intuitive understanding of justification come together. This is what it is to argue for a truth norm on the basis of its explanatory role within epistemology.

I'll argue that a truth norm can play this explanatory role in the thesis as a whole. But I first want to show that this way of arguing for a truth norm is a good way to
arguing for something, and in particular a good way to argue for a norm. I’ll argue for this, in §6.1, with an analogy with practical philosophy. I’ll then end this chapter, in §6.2, by discussing an inherent limitation of arguing for a truth norm in this way.

6.1. The appeal to this methodology in ethics

It’s plausible to assume, I am suggesting, that if a particular norm accounts for our normative intuitions, that by itself provides support for that norm. By ‘support’ I only mean first-order normative support, i.e. if a norm on belief or action accounts best accounts for our normative intuitions, I only think that provides support for that norm over other alternative norms on belief or action. I don’t think, on the other hand, that a norm accounting for our normative intuitions can by itself provide itself metanormative normative conclusions about the source or ground of moral or epistemic normativity, though I’ll discuss this briefly in §6.2.

We can see how accounting for normative intuitions provides support for a norm through an analogy with practical philosophy. When discussing what we morally ought to do, it is plausible to assume, I’ll suggest, that if a moral norm accounts for our intuitive understanding of what we morally ought to do, that by itself provides support for that norm. If this methodology is acceptable in ethics, I see no reason why we can’t appeal to its analogue in epistemology, i.e. to argue for a norm on belief on the basis of how it accounts for our intuitive understanding epistemic justification.

I don’t pretend, however, to be giving give a knockdown argument in defence of

---

5 This methodology for arguing for a truth norm may seem to some to be Rawls’s methodology of ‘reflective equilibrium’. I agree that it is very similar. Arguing for a truth norm on the basis that it gives prescriptions that cohere with our intuitive understanding of when someone is justified looks like what Rawls talks about when he speaks of ‘testing’ our moral and political theories against our ‘considered judgments’ about particular cases (1999, xii, 18, 42). And my claim that a truth norm can explain the more abstract semi-theoretical aspects of our intuitive understanding of justification recalls Rawls’ claim that moral theorising has to appeal to both ‘contingent assumptions’ and ‘general facts’ (1999, 44).

Rawls does not, however, view this methodology to be particularly original. He claims that it is essentially the same methodology used by Aristotle, and by most of “most classical British writers through Sidgwick” (Rawls 1999, 45). I think Rawls is right about this, and I think this means we don’t need to discuss the minutiae of Rawls’ particular understanding of this methodology in order to use it. We can appeal to it simply because it is, as I’ll show, overwhelmingly the predominant way of arguing for (or against) particular norms.
this methodology. Doing that would take us too far afield.\(^6\) I only take myself to be giving a *prima facie* case in favour of there being an alternative way of arguing for a truth norm, a way that does *not* involve arguing that it is constitutive of belief. Given how much the idea of a truth norm has been run together with the idea that belief is *constitutively* subject to such a norm, it seems worthwhile to consider an alternative way of arguing for a truth norm that does not run these ideas together.

I'll illustrate this point with regards to, in particular, discussions of *consequentialist* theories in normative ethics. Two aspects of these discussions support, as I'll show, the assumption that one can provide support for a moral norm by showing that it accounts *for* our normative intuitions. Firstly, this assumption is supported by the fact that it in implicit in many of the most prominent arguments for and against consequentialism. Secondly, I'll use the example of consequentialism to outline how implausible it is to deny this assumption – i.e. to claim that one *must* provide support for a moral norm that is somehow independent of how it accounts for our moral intuitions.

Let's first consider arguments for or against particular consequentialist theories. Consider the following simple maximizing consequentialist norm about what one *morally* ought to do:

```
MAXIMIZE: One ought to perform the action which, out of all the available alternatives, has the best consequences.
```

In order to defend this norm, would we need to provide reason that is *independent* from our intuitive understanding of what one morally ought to do? This is certainly not how discussions of consequentialist norms like MAXIMIZE typically go. Rather, it seems to be assumed, by both defenders and critics alike, that *if* MAXIMIZE accounted for all (or even many) of the elements of our intuitive understanding of what one morally ought to do, that would *by itself* provide support for MAXIMIZE.

This assumption is manifest in many of the most prominent arguments for and against consequentialism. Let’s start with the arguments in favour. Such arguments often claim that a consequentialist norm like MAXIMIZE is uniquely able to deal with one or more elements of our intuitive understanding of what we morally ought to do.

\(^6\) For much more detailed defence of a methodology of roughly this kind, especially in relation to ethics, see the essays in Daniels 1996.
Furthermore, just as I indicated with regards to a truth norm, the elements of our intuitive understanding that MAXIMIZE is claimed to explain can be more concrete or more abstract.

On the more concrete end, some argue that a consequentialist norm like MAXIMIZE is the only moral norm that can explain particular things one ought to do. For example, some argue that only a consequentialist norm like MAXIMIZE can explain why we have duties to animals (Bentham 1789, 282, fn.1), or why we have duties to future generations (Parfit 1984, Ch. 16). Getting slightly more abstract, some argue for maximize because it provides a unifying explanation of how all of particular intuitive or common-sense moral judgements hang together (see e.g. Sidgwick 1907, Bk. VI, Ch. III). At the much more abstract end, we might motivate MAXIMIZE by the more theoretical idea – which Philippa Foot claims is at the heart of philosophers’ attraction to consequentialism – “that it can never be right to prefer a worse state of affairs to a better” (1985, 198). Alternatively, some claim that MAXIMIZE finds support in the equally theoretical idea that what one morally ought to do essentially is what one ought to do from an impartial or universal perspective (see e.g. Singer 1993, 8–15). All of these are arguments that we should adopt a consequentialist norm like MAXIMIZE because it can explain some element of our intuitive understanding of morality.

Furthermore, this assumption is not only present in arguments for MAXIMIZE, but also in arguments against it. This can be seen by the fact that a number of the most prominent criticisms of a norm like MAXIMIZE are, in effect, claims that MAXIMIZE fails to do justice to some aspect of intuitive morality.

This is true of many of the most prominent criticisms of consequentialism. It is true of criticisms that appeal to the fact that consequentialist norms like MAXIMIZE prescribe unconscionable trade-offs that look prima facie morally wrong because they involve violating someone’s rights. Such cases include the trolley problem (Foot 1967, 23), the case of a transplant surgeon who can save five people’s lives by killing one healthy person and taking his organs (Thomson 1976, 206), or the case in which the authorities can prevent riots and lynchings by punishing an innocent person for a crime (McCloskey 1957, 468–69). It is also true of Bernard Williams’ examples of Jim and the Indians and George the Chemist (in Smart and Williams 1973, 93–100), which he claims show that a consequentialist norm like MAXIMIZE fails to account for the value of integrity. It is also true of the claim that a consequentialist norm like MAXIMIZE is overdemanding, because
it cannot account for why one ought to carry out any non-moral personal projects (Wolf 1982).

Each of these criticisms of MAXIMIZE essentially claims that we should reject it because there is some particular element of our intuitive understanding of what we morally ought to do – i.e. rights, integrity, personal projects – which MAXIMIZE fails to account for.

Furthermore, this assumption is also shared by those who defend consequentialism in response to these arguments. This is true of those who claim that MAXIMIZE does not in fact fail to account for one or more of these elements of intuitive morality (see e.g. Railton 1984; Jackson 1991). This is also true of those who give modified – though still recognisably consequentialist – norms that do not have the counterintuitive prescriptions that MAXIMIZE does. This includes both rule-consequentialists (see e.g. Brandt 1963; Hooker 2000), as well as consequentialists who appeal to agent-relative values in order to get over these problems (Broome 1991, 1–6). This is also true of those who opt for a non-consequentialist moral theory in response to the above problems for consequentialism (Kamm 2000). All of these are attempts to defend norms that do succeed in accounting for our intuitive understanding of what we morally ought to do.

This survey of arguments for and against consequentialism – and responses to these arguments – shows that much of what it is to argue for a moral norm is to argue that it accounts for our intuitive understanding of what we morally ought to do. Arguments that are independent of our moral intuitions are not typically seen as needed. This gives us initial reason to think that this is a good way of arguing for a norm, and that it is a methodology we can carry over to epistemology.

We can further strengthen this point by highlighting how implausible it would be to deny the assumption that one can argue for a norm by showing that it accounts for our normative intuitions. Again I’ll make the point by analogy with moral norms.

Firstly, the implausibility of denying this assumption can be highlighted by the way one might reject the arguments for MAXIMIZE that I listed above. The way one would reject these arguments would typically involve claiming either a) that the element purportedly explained by MAXIMIZE is not in fact part of our intuitive understanding of morality, b) that other alternative moral norms can explain it, or c) that MAXIMIZE in fact fails to explain it. It seems much less plausible to argue against these arguments by claiming that even if MAXIMIZE did account for our moral intuitions that wouldn’t
provide support for it.

The same seems to me to be the case in epistemology. It doesn’t seem plausible to admit that a truth norm does account for our intuitive understanding of justification, but deny that provides any support for a truth norm.

Secondly, denying this assumption – and therefore not appealing to our intuitive understanding of what we morally ought to do – looks like it provides far too meagre a basis to argue for any moral norms. It would seem difficult to argue for any substantive moral theory if we cannot appeal to our intuitive understanding of what we morally ought to do. This is because not appealing to how well a moral theory accounts for our moral intuitions looks like it leaves things far too open.

While attempts have been made to argue for consequentialism (and other moral theories) without appealing to moral intuitions, I think we should be sceptical about whether these attempts can succeed.

One example of such an attempt is G. E. Moore’s claim that a consequentialist account of rightness is analytic. He claims that “the assertions ‘This action is right’ or ‘is my duty’ are equivalent to the assertion that the total results of the action in question will be the best possible” (Moore 1903, 47; cf. 196-98). Another example is Richard Brandt, who claims to argue for rule-consequentialism purely on the basis that it is the rational moral theory to adopt. By ‘rational’ Brandt means “to refer to actions, desires, or moral systems which survive maximal criticism and correction by facts and logic” (1979, 10)

While I don’t have space to go into it in detail, I want to suggest that we should be sceptical about opting for one of these lines of argument. Firstly, it is implausible to argue, as Moore does, that one moral theory is analytic, since that has the implausible result that those who put forward alternative moral theories are saying something not just false, but contradictory. Secondly, appealing to rationality alone, as Brandt claims to do, doesn’t look like it’s going to support one substantive moral theory over another. In particular, abstract principles of logic, rationality, or theory choice look far too ‘thin’ and formal to enable us to decide between, say, a norm like MAXIMIZE and other alternative consequentialist or deontological moral theories. It looks very likely that ‘purely rational’ considerations are going to leave this an open question.

Again, the same seems to be the case in epistemology. If we don’t appeal to our intuitive understanding of justification, that’s likely to leave open the very question we want to answer, i.e. whether we should favour a truth norm over a knowledge norm, or
over a purely evidential norm. Furthermore, as I’ve shown, a number of arguments for a truth norm on the basis of its purported explanatory role in the philosophy of mind do in fact leave it open whether belief is subject to specifically a truth norm.

6.2. A limitation of this methodology

Although I think we can argue for a truth norm purely on the basis of its explanatory role in epistemology, I just want to flag a way in which this makes my conclusion essentially limited. I don’t think, however, that this limitations of my methodology is problematic.

The limitation in arguing for a truth norm because it accounts for our intuitive understanding of justification is that our conclusion has to remain a strictly first-order normative conclusion. What I mean by this is that I don’t think we can make any metanormative conclusions about the source of epistemic normativity arguing in the way I’m suggesting.

This is essentially because I do not - whereas many the other philosophers I’ve discussed do – argue that it is constitutive of the attitude of belief that it is subject to a truth norm. Showing that belief is constitutively subject to a truth norm has the prospect of explaining the source of epistemic normativity in a similar way that some moral philosophers – especially some Kantians – claim that moral normativity comes from what is constitutive of agency (Rawls 1980; O’Neill 1989; see e.g. Korsgaard 1996, Ch. 3-4; 2009). Following Rawls, this is typically known as a ‘constructivist’ metaethics. Some defenders of a truth norm explicitly take themselves to be giving a constructivist explanation of the source of epistemic normativity (Wedgwood 2007; Evans and Shah 2012). I suspect the reason why so many defenders of a truth norm argue that it is constitutive of belief is because they want to make both first-order normative and metanormative conclusions, i.e. to conclude both that one ought to believe something if and only if it’s true and that one should do so because that’s constitutive of belief.

I take it that the methodology I’ve outlined, however, can only show that belief is subject to a truth norm, but not that it is essentially so. Therefore, it does not have the same prospects at explaining the source of epistemic normativity.

I don’t think, however, that this limitation of my methodology is problematic. This is because, firstly, a constructivist story about the source of epistemic normativity is not the only game in town. All we need is some explanation of the source of epistemic
normativity, and, in particular, one which allows us to retain the claim that our intuitive understanding of when someone counts as justified is a reliable guide to what one ought to believe. There are a variety of possible explanations about the source of epistemic normativity – both realist and anti-realist – that would have this result, so we shouldn’t view defending a truth norm as essentially tied to a constructivist explanation.

Secondly, there are notable problems a constructivist account of moral normativity has to answer, problems that would carry over to a constructivist account of epistemic normativity. A prominent problem for constructivism about moral normativity is raised by David Enoch. Enoch points out that even if it’s true that it is constitutive of agency that one ought to φ, it still seems that I can ask whether I ought to be an agent, rather than a ‘schmagent’, i.e. something for which it’s not the case that it ought to φ, but in other respects is like an agent (Enoch 2006). This seems to indicate that there is a kind of normativity which cannot be explained by what is constitutive of agency. And that normativity looks like a better candidate to be genuinely moral normativity, because it looks more categorical than whatever ‘oughts’ are constitutive of agency.

If someone were to offer a constructivist account of the source of epistemic normativity, such a view would seem to face much the same problem. That is to say that if it is constitutive of belief that one ought to believe the truth, there is still an open question about whether I ought to be a believer rather than a schmeliever (Papineau 2013, 75–79). Therefore, even if a truth norm is constitutive of belief, we would have to answer an objection of this kind before such a truth norm could play any metanormative role in explaining the source of epistemic normativity. Therefore, although my way of arguing for a truth norm – which purports to show that belief is subject to a truth norm, but not that is essentially so – does not promise to answer the metanormative question about the source of epistemic normativity, that limitation is unproblematic. That is because it’s not straightforward that belief being essentially subject to a truth norm would answer this metanormative question either.

7. Conclusion

In summary, we should be happy to argue for a norm on the basis of how well that norm accounts for our first-order normative intuitions. The example of arguing for moral norms both showed how commonplace this method is, and how difficult it would be to argue for any particular norm without appealing to this method. Although I’ve only
discussed the one example of moral norms, I suspect that much the same is the case in any area of philosophy that involves arguments for or against particular norms, such as political philosophy, aesthetics, or decision-theory.

To repeat, I don’t want to suggest that I’ve provided a knockdown argument in favour of this way of arguing for a norm. However, I hope I’ve done enough to show that this is a plausible alternative way to argue for a truth norm. As I’ve said the rest of the thesis will provide the argument that a truth norm in fact does provide a unifying explanation of justification. But even if I’m wrong about that, the conclusion of this chapter would still hold. We should still recognise that a truth norm does not stand and fall with the success or failure of the psychological arguments that a truth norm is constitutive of belief. We do not need to argue that a truth norm is constitutive of belief or appeal to its explanatory role in the philosophy of mind in order to use a truth norm in epistemology. There is an alternative way of arguing for a truth norm from within epistemology, and the rest of this thesis outlines this alternative.
This chapter gives an epistemological argument – of the kind I suggested in the last chapter – for the claim that belief is subject to a truth norm. I argue that a truth norm can explain one central feature of justification, how justification is unified, i.e. the fact that all justified beliefs have something in common in virtue of which they count as justified. If there was nothing in virtue of which all justified beliefs counted as justified, then justification would be, I argue, problematically heterogeneous. If belief is subject to a truth norm, however, then the problematic heterogeneity of justification is averted, because we have an easy answer to the question of what all instances of justification have in common: anything that justifies a belief counts in favour of its truth.

I’ll proceed as follows. In §§1-2, I’ll argue that there must be something in virtue of which all justified beliefs count as justified, because if there were not, there would be implausible consequences.

I do this by discussing two different ways in which justification looks prima facie heterogeneous. In §1, I discuss the various different methods which can justify belief. In §2, I discuss the various different properties that bear on whether a belief is justified, which are sometimes referred to as the different ‘theoretical virtues’, such as consistency, simplicity, etc. I argue that in both cases it would be problematic if there were not something which unified these different methods and properties, and I argue that a truth norm can play this unifying role.

After this, in §3, I discuss alternative possible explanations of the unity of justification; in particular, I discuss whether a knowledge norm or an evidence norm on belief can play this unifying role. I’ll argue that these alternatives either fail to provide a unifying explanation of justification, or they have other drawbacks when compared to a truth norm.

1. Different methods that justify beliefs

The claim of this section is that a truth norm can unify the variety of methods that justify beliefs. I’ll do this first by arguing, in §1.1, that there needs to be something that unifies these various methods, because it would be problematic if the different methods that
justify beliefs were what I'll call ‘radically heterogeneous’, i.e. if there were nothing that the variety of methods had in common in virtue of which they result in justified beliefs. I'll then explain, in §1.2, what kind of thing could unify justification, by discussing some clearly inadequate ways to unify it. I'll then explain, in §1.3, how a truth norm can play this unifying role.

1.1. Why something must unify different methods that justify beliefs

For justification to be unified, all justified beliefs must have something in common in virtue of which they count as justified. This is not to say that justified beliefs must be exactly the same in every respect, just that, at some level of description, all justified beliefs count as justified for the same reason. I'll outline what this claim means – and go on to to defend it – first by thinking about the different methods that justify beliefs.

A belief’s justification seems to have something to do with the method by which that belief is formed. Take the following belief:

MATHEMATICAL: the belief that 43-17=26.

What determines whether someone is justified or unjustified in holding this belief? This seems to come down (in part) to the method by which one came to have the belief. If someone believes it on the basis of carefully carried out mental arithmetic or having used a calculator, then the belief seems justified. On the other hand, if someone believes it on the basis of sloppy mental arithmetic which luckily produces the right answer, or on the basis of a guess, or from reading it in the tea leaves, then the belief doesn’t seem justified.

But is there anything about the method by which beliefs are formed that can explain what all justified beliefs have in common, and what makes them different from unjustified beliefs? This can look difficult to answer, given the differences between the methods by which we form beliefs about different subject matters. Consider the following beliefs:

MATHEMATICAL: the belief that 43-17=26.
PERCEPTUAL: the belief that it’s raining outside (after looking out the window).
LOGICAL: the belief that there are no true contradictions.
ECONOMIC: the belief that the UK economy will go into recession in the next year.
PSYCHOLOGICAL: the belief that Christopher is going to the swimming pool (after
seeing him grab his trunks and goggles on the way out).

When these beliefs are justified, it seems that this will have something to do with the method by which they are formed. However, the methods that result in such beliefs themselves look very different. Mental arithmetic, for example, looks like a very different method from whatever method I use to form the belief that Christopher is going to the swimming pool. In this way, it seems difficult to find one thing that all these methods have in common that makes for justification across the board. Furthermore, this difficulty is only going to grow if we consider the full range of different methods that can lead to justified beliefs.

This variety among the different methods that justify beliefs may lead one to conclude that there is no one thing that all justified beliefs have in common in virtue of which they count as justified. This is to think of justification as what I’m calling ‘radically heterogeneous’.

A claim like this can come in more and less extreme versions. A more extreme version would claim that ‘justified’ is an ambiguous term, one which somewhat arbitrarily applies to beliefs produced by different methods, much like how the term ‘bank’ arbitrarily applies to both financial banks and river banks. A less extreme version would claim that ‘justified’ is a family resemblance concept, i.e. that although each individual justified belief has something (and perhaps much) in common with some beliefs justified by a distinct method, there is no one thing that beliefs justified by different methods all have in common. Both of these options count as thinking of justification as radically heterogeneous in my sense because both involve denying that there is one thing that all justified beliefs have in common in virtue of which they count as justified.

I want suggest that this would be a mistake. The conclusion that justification is radically heterogeneous has implausible consequences. This is because, despite their variety, the different methods that justify belief interact in ways that would be inexplicable if justification were radically heterogeneous.

For example, it seems that a belief formed by one method can have its justification defeated by a completely different method. For example, say I come to believe that the hypotenuse of a particular right-angled triangle is 10cm using Pythagoras’s theorem and mental arithmetic. However, it also happens that the Fields Medalist Maryam Mirzakhani is looking over my work and tells me that the hypotenuse is 9cm long. Given that I trust her mathematical skills over mine, I revise my original belief. In
this case, my belief that the hypotenuse is 10cm – which was the result of mental arithmetic – is defeated by a completely different method – testimony.

Although this is just one example, it is commonplace for different belief-forming methods to interact in this way. For example, most ways we have of ‘checking’ whether one belief-forming method in fact leads to justified beliefs involves appealing to a different method.

If justification were radically heterogeneous in either of the ways I mentioned earlier, then it would be completely inexplicable how different methods could interact in this way.

This is especially clear if ‘justified’ is supposed to be an ambiguous term, which only arbitrarily refers to beliefs formed by different methods. If that were the case, then we have no explanation as to why the testimony of the Fields Medalist is not entirely irrelevant to the deliverances of my mental arithmetic. Since two methods clearly do interact in this case, as in many others, we can dismiss the idea that ‘justified’ is an ambiguous term.

The claim that justification is a family resemblance concept also, I suggest, cannot provide a good explanation as to how different methods that justify beliefs interact. This is because justification being a family resemblance concept entails the possibility of cases in which one justified belief formed by one method and another justified belief formed by another method have nothing in common. In such a case, the family resemblance view would make it just as inexplicable how those different methods could interact as the ambiguity view does. This means that a family resemblance view of justification will possibly face the same issues as a simple ambiguity view about how different belief-forming methods interact (though without a specific family resemblance view on the table it’s difficult to provide a particular example to show this as we could with the ambiguity view). Therefore, given how commonplace it is for belief-forming methods to interact, it seems best to avoid a family resemblance view of justification as well.

By this I don’t mean to suggest that there is anything problematic about the very idea of a family resemblance concept. I’m happy to grant that sometimes species are tied together into a genus by relations of family resemblance; plausibly, the concept of a game works like this (Wittgenstein 1953, para. 67). All I am claiming is that justification cannot be a family resemblance concept, because that wouldn’t be able to account for how different belief forming methods interact in determining whether a belief is justified.
It is a mistake, therefore, to react to the variety among the different methods that justify beliefs by concluding that justification is radically heterogeneous. To explain the interaction between methods, we need to think of all justified beliefs as having something in common in virtue of which they count as justified.

1.2. What could unify different methods that justify beliefs?

Saying what unifies the different methods that justify beliefs involves a tricky balancing act. On the one hand, the explanation can’t be too specific, i.e. so tied to a particular method that it ignores the diversity of methods that can lead to justified beliefs. On the other hand, it can’t be too minimal, i.e. it can’t be a feature that all methods that justify beliefs only trivially share, because such a feature wouldn’t be able to do the requisite explanatory work of explaining why different methods interact.

I’ll suggest, in §1.3, that a truth norm can strike the right balance here. But we can outline these two desiderata of a) not being too specific and b) not being too minimal by highlighting accounts of justification which fail to meet each of them.

What’s the problem if the proposed common feature of methods that lead to justified beliefs is too specific? What I have in mind is the suggestion that one particular method makes for justification in all cases. This suggestion seems doomed to failure because there is such a variety of methods by which justified beliefs are formed, and no one of those methods seems applicable to all justified beliefs. Trying to show that there is in fact just one method of coming to justified beliefs is bound to ignore the diversity in justification that is really there.

We can illustrate this with the example of Early Modern rationalism and empiricism. The versions of these views I will outline are perhaps historically inaccurate, but if so, this will not matter for the illustration.

Rationalism and empiricism each take a particular method involved in a paradigm case of justification to be what makes for justification in all cases. Because of this, they are often criticized for failing to account for justification in cases that involve methods different from those involves in the paradigm cases.

For Descartes, on the rationalist side, the paradigm case of justified belief was one which one could guarantee to be true just by thinking about its content – in Descartes’ terms, an idea of which one could have a clear and distinct perception of its truth (1641, AT VII: 35). Such a method looks too specific to be one we genuinely use
whenever we form justified beliefs. In particular, it looks difficult to see how such a method could explain why we are justified in cases in which we don’t have this kind of guarantee, such as cases of fallible empirical beliefs about the external world. Of course, Descartes himself thought that clear and distinct perception could ground the justification of our empirical beliefs. This because he thought he could prove, through clear and distinct perception, the existence of a non-deceiving God who would guarantee that our empirical beliefs about the external world are not systematically false (Descartes 1641, AT VII: 70-71, 87-91). But if we are more sceptical than Descartes was that we can prove the existence of God through reason alone, it looks inevitable that clear and distinct perception will be too specific a method to explain how all our justified beliefs count as such.

On the empiricist side, Hume took the association of subjective experiences to be the paradigm of justification (1739, 1–7), and the consequence of this was that he was forced into scepticism about a number of things which we don’t seem to know through the association of subjective experiences. For example, one result was that Hume could not give any account of how we are justified in thinking that there is a self or anything that unifies conscious experience (1739, 251–54). Hume himself worried about this consequence of his epistemology (1739, 633–36). Many others – including perhaps most notably Kant (see e.g. 1781/87, A 106-8, B 131-36) – have taken Hume’s failure to account for our knowledge of the self to be problematic.

A common flaw in both Descartes’s and Hume’s views seems to be that they took the mark of justification to be something too specific, which led to them having trouble dealing with cases of justification that were unlike their paradigm case. It looks highly likely that any claim that a specific method makes for justification across the board will have the same flaw, and will fail to do justice to the genuine diversity among methods that justify beliefs.

On the other hand, the common feature of different methods that explains why they all lead to justified beliefs can’t be too minimal, i.e. it can’t be a feature that all different methods that lead to justified beliefs trivially share. If all methods that lead to justified beliefs trivially share this feature, then it doesn’t seem that the feature can explain why beliefs produced by these different methods all count as justified.

For example, for all methods that lead to justified beliefs, it seems true that one has reason to believe that p if one’s belief that p was produced by one of those methods. It
also seems true that one would be rational if one believed that p if one’s belief that p was produced by one of those methods. But these are features that all justifying methods trivially share, and it does not seem that a feature that justifying methods trivially share can do any explanatory work of explaining why a particular method is a justifying one.

The fact that a feature justifying methods trivially share cannot do this explanatory work can be seen if we return to my earlier example of Maryam Mirzakhani correcting my mental arithmetic. A feature that justifying methods trivially share has no prospect of explaining why one method can bear on the justification of another belief produced by a completely different method. Saying that both mental arithmetic and testimony provide reasons for belief doesn’t clear up how my belief that the hypotenuse is 10cm, which resulted from mental arithmetic, can be defeated by the completely different method of testimony. I then just want to know why both mental arithmetic and testimony can provide reasons for belief, and how they are supposed to interact. That is just as much a mystery as the question of how these very different methods can produce justified beliefs.

Velleman makes a similar point to this in his discussion of ‘formal’ and ‘substantial’ aims. A merely formal aim of a practice, in Velleman’s terms, is one that is defined solely as that which that practice aims at (1996, 176–68). To illustrate, the formal aim of chess is winning, whereas the substantial aim is checkmating the opponent. A formal aim of a practice is uninformative, whereas a substantial aim is not.

In Velleman’s terms, being based on reasons is a purely formal aim of belief, because it is trivially the case that if one is justified in believing that p one has reason to believe that p. It will, therefore, be necessarily uninformative, and won’t be able to explain why a particular method is a justifying one.¹

One may object to my point here by claiming that a justified belief is not necessarily a belief one has reason to have. This would be the case if we gave ‘justified’ an internalist reading and ‘reason to believe’ an externalist reading.

This is a possible view, but it doesn’t touch the point I’m making. This is because it has the result that it’s simply false that all methods that justify are methods that give

¹ Velleman also argues, as I will do, that truth is the substantial aim of belief (1996, 180–82). For more discussion of Velleman’s distinction between formal and substantial aims, and its application to the case of belief, see Cowie 2013.
one reason for belief. So it then can’t explain what all methods that justify have in common. The same will be the case if we give ‘justified’ an externalist meaning and ‘reason to believe’ an internalist one.

All methods that justify will be methods that give one reason to believe only if we give both ‘justified’ and ‘reason to believe’ the same sense, whether internalist or externalist. If they have the same sense, then all methods that justify will trivially be methods that give one reason to believe, and therefore appealing to this fact cannot, for the above reasons, explain why a particular method is a justifying one.

In summary, what unifies different methods that justify beliefs can neither be too specific nor too minimal. Choosing a particular method as a paradigm case of justification will typically be too specific. Appealing to other explicitly epistemological concepts like reasons or rationality will – if they are shared by all justified beliefs – typically be too minimal.

1.3. How a truth norm unifies different methods that justify beliefs

Belief being subject to a truth norm can, I suggest, provide the right kind of explanation of what unifies different methods that justify beliefs. I’ll illustrate with the following simple truth norm (though I’ll end up, in Chapter 5, defending an alternative formulation):

**OUGHT**: One ought to believe that p if and only if p is true.

A truth norm like this can, I’ll argue, explain what unifies justification. That is not to say that justified beliefs are all and only those beliefs that are true. That would be a bad suggestion as to what unifies justified beliefs, since it’s false. There are unjustified true beliefs and justified false beliefs. Rather, the way a truth norm like OUGHT can unify justification can be seen by the way, which I outlined in Chapter 1, in which a truth norm is supposed to explain justification. OUGHT is supposed to explain justification because one cannot follow it directly, i.e. one can’t ‘just tell’ whether one’s beliefs are true. In this sense, OUGHT is an ‘objective’ norm, because of this, one is going to have to take some means to the end of conforming to it. This is going to have to involve following some other ‘subjective’ norms – i.e. norms that one can follow directly – as means to the end of conforming to OUGHT. Following the subjective norms one should follow as a means to the end of conforming to OUGHT, the account claims, will just involve having beliefs
that we intuitively take to be justified.

In this way, it is through these subjective norms that OUGHT explains justification. But because these subjective norms are all orientated to the same end – i.e. the end of believing that p if and only if p is true – OUGHT also provides an answer to the question of what unifies the subjective norms, and thereby to the question of what unifies justification. Roughly, what unifies all the different kinds of justification is that they all count in favour of the truth of the proposition believed, which we can express as follows:

**TRUTH CONNECTION:** A factor F justifies a belief that p if and only if F counts in favour of p being true.

As I outlined above, TRUTH CONNECTION follows from OUGHT, given the way in which such a truth norm explains justification.

Furthermore, TRUTH CONNECTION can provide the right kind of explanation of what unifies the different methods that justified belief, i.e. one that is neither too specific nor too minimal. It is a non-trivial future which is shared by all the different methods that justify beliefs, but it is liberal enough to allow us to recognize the genuine diversity among these methods.

Before I say how TRUTH CONNECTION can unify justification, I want to make a few points about it. I first want to stress the limits of the conclusions of this chapter in two respects.

Firstly, TRUTH CONNECTION is not intended to be the whole story about how a truth norm is supposed to explains justification. It leaves a number of questions still open. For example, it leaves open the question of how much justification a belief needs in order to be justified. This is a question we’ll touch on briefly at the end of this chapter and in Chapter 6 I’ll argue that a truth norm can answer this question. However, although TRUTH CONNECTION is too weak to answer that question, it is strong enough to answer the question of this chapter, i.e. to explain what unifies justification.

Secondly, although I’ve claimed it follows from a truth norm, I don’t think that TRUTH CONNECTION can only be appealed to by those who defend a truth norm. I think that philosophers who claim that there is some essential connection between justification and truth can appeal to TRUTH CONNECTION to explain what unifies justification. For example, I think reliabilists about justification can appeal to a claim like TRUTH
CONNECTION to explain what unifies justification (see e.g. Goldman 2001). In this way, the fact that a truth norm unifies justification doesn’t provide reason to favour a truth norm approach to justification over other versions of the claim that there is an essential link between justification and truth (such as reliabilism).

I also want to make two points about TRUTH CONNECTION’s phrasing.

Firstly, I’ve chosen to speak of ‘factors’ justifying beliefs because at this stage I want to be maximally inclusive about what can justify beliefs. I want to state the truth connection in a way that would allow beliefs to be justified by other beliefs, perceptual experiences, propositions, or facts. I don’t want any particular view about what can justify beliefs to come out of TRUTH CONNECTION.

Secondly, I want to be silent on whether ‘counts in favour’ should be read as counting in favour objectively, or only from the subject’s perspective. This is because, as I indicated above, I want to leave open the possibility that both externalist and internalist accounts of justification can agree with TRUTH CONNECTION.

But even while remaining neutral on these questions, it seems that TRUTH CONNECTION can explain what unifies justification, i.e. it can tell us what different methods that lead to justified beliefs have in common.

According to TRUTH CONNECTION, whenever a belief that \( p \) is justified, it is justified by factors that count in favour of the truth of \( p \). This means that a method will justify if beliefs produced by that method are likely to be true. Again, as with ‘count in favour’, I want to be neutral about whether we interpret ‘likely to be true’ as objectively likely or likely from the subject’s perspective.

This can provide an explanation of what all methods that lead to justified beliefs have in common. The fact that my belief that \( 43-17=26 \) was the result of careful mental arithmetic counts in favour of that belief being true. If I had formed that belief as the result of a lucky guess, that would not count in favour of that belief being true. The fact that I have formed my belief that Christopher is going to the swimming pool on the basis of what I take him to want and believe counts in favour of the truth of that belief. If I had just assumed that Christopher was going swimming because that would be why I would be heading into town, ignoring the fact that he may have different desires to me, that would not count in favour of that belief being true.

Furthermore, TRUTH CONNECTION can explain what unifies different methods that can provide justification for belief, without ignoring the diversity among different
methods. This is because there are many different ways of figuring out whether something is true. There is a massive variety of factors that can count in favour of the truth of a belief, and different kinds of factor will count in favour of the truth of beliefs about different kinds of subject matter.

Because of this, we should expect different methods to justify beliefs about different subject matters. Providing a proof for mathematical or logical propositions is a factor that counts in favour of those propositions being true, and therefore, according to my analysis, justifies believing those propositions. But providing a similar kind of proof for an economical or psychological proposition looks out of place. This is because different kinds of factors count in favour of the truth of economical or psychological propositions. But we can still say that what justifies beliefs about economics and psychology has something in common with what justifies beliefs about logic and mathematics, because all are factors that count in favour of truth. In this way, we explain what unifies justification, while respecting the diversity of different methods.

Because of this, the fact that different methods that lead to justified beliefs look heterogeneous in many respects should no longer lead to any worries about ‘justified’ being ambiguous. This is because despite their heterogeneity, there is one feature non-trivially shared by all methods that lead to justified beliefs: they count in favour of the truth of the beliefs they produce.

2. Different properties that bear on the justification of a belief

Having explained how TRUTH CONNECTION explains what unifies diverse methods that justify beliefs, I now am going to suggest that we can give the same explanation of what unifies diverse properties that bear on the justification of belief.

When thinking about whether a belief is justified, it seems many different properties of that belief bear on that question. The belief might be said to be ad hoc, and because of that lacking in justification. The belief might be the simplest thing to believe in the circumstances, and because of that appear to be justified. What do all these properties have in common in virtue of which they bear on justification? Again, I will argue that a truth norm can explain this.

I’ll do this by first, in §2.1, outlining the different properties that bear on the justification of belief. I’ll then argue, in §2.2, that it would again be problematic if there were nothing all these different properties had in common. I’ll then argue, in §2.3, that a
truth norm – through my formula TRUTH CONNECTION – can explain what unifies these different properties.

2.1. The variety of properties of a belief that can bear on its justification

In speaking of different ‘properties’ of that bear on a belief’s justification, I am referring to the variety of properties of a belief – such as simplicity and consistency – which we might appeal to when considering whether a belief is justified. These properties are most often discussed in the philosophy of science as the different ‘theoretical virtues’ or ‘criteria of theory choice’ used to assess different scientific theories. The discussion of these different properties in the philosophy of science will provide the springboard for my argument in this section.

The reason I’m going to use these examples taken from the philosophy of science is because many of the so-called theoretical virtues look like they bear on the justification of belief not just in scientific contexts, but across the board. It seems we can justify both a scientific theory and a belief about a particular matter of fact by appealing to simplicity. It also seems that both a scientific theory or a belief about a particular matter of fact can be unjustified because they are ad hoc. Many – but not, as we’ll see, all – of the theoretical virtues also bear on justification across the board. Because of this, I will later on speak more often of particular scientific ‘hypotheses’ than ‘theories’, because they are more comparable to particular beliefs.

The theoretical virtues are discussed in relation to the question of how scientists choose the best theory to explain a particular phenomenon. This is not simply about which theory is best supported by the evidence, because all the evidence there is can equally support many theories.

To illustrate, a geocentric astronomical theory can be equally well supported by the evidence as a heliocentric one, if enough compensatory adjustments are made to the geocentric theory to deal with recalcitrant evidence. Furthermore, it seems that one theory can be preferred to another theory in some cases even if it is less well supported by the evidence than another. Kuhn suggests this to have been the case in Copernicus’s heliocentric astronomy: “Copernicus’s system … was not more accurate than Ptolemy’s until drastically revised by Kepler more than sixty years after Copernicus's death” (Kuhn 1977, 323). Because of these considerations, it is suggested that scientists in fact decide between different theories by appealing to theoretical virtues other than how well the
theory is supported by the evidence.

Philosophers of science don’t all agree about what the theoretical virtues are. Representative lists are given by Hempel (1966, chap. 4), Kuhn (1977, 321–24), and Quine and Ullian (1978, chap. 6). I’ll outline some of the most commonly cited theoretical virtues.

One theoretical virtue is conservatism (Quine and Ullian 1978, 66–67). A theory or hypothesis is more conservative than another if it conflicts with less of our existing beliefs. Another way of putting this point is to say that a theory or hypothesis that is more consistent with our background beliefs is preferable to a theory or hypothesis that is not (Kuhn 1977, 323). It also seems fairly clear that how conservative a belief is can bear on its justification across the board and not just in a scientific context. It seems that I can be more justified in believing that p, rather than an alternative hypothesis, q, because p conflicts with fewer of my other beliefs than q.

Another virtue is modesty (Quine and Ullian 1978, 68–69). A theory or hypothesis is more modest than another if it is logically weaker, i.e. it entails less than the other theory or hypothesis. This also seems to bear on justification across the board. This fact is clear if we focus on a particular way of showing a belief to be unjustified. A common way of showing that a belief is unjustified is to point out that the reasons offered to support that belief only seem to support a much weaker belief.

Another commonly mentioned virtue is simplicity (Popper 1959, chap. 7; Kuhn 1977, 324; Quine and Ullian 1978, 69–73; Sober 2015). Often a definition is not given, and many disagree about a) exactly what the simplicity of a theory or hypothesis amounts to and b) why simplicity is a theoretical virtue (see Fitzpatrick 2016, secs. 3–4 for a summary of this disagreement). However, it is relatively uncontroversial that simplicity is a theoretical virtue, and we can think of a simple scientific hypothesis or theory as one that does not multiply assumptions beyond necessity.²

A commonly given example illustrating the virtue of simplicity is how Copernicus’ heliocentric astronomy improved over its geocentric predecessors. Hempel

---

² This phrasing is close to the classic medieval formulation of Ockham’s razor - ‘entia non sunt multiplicanda praeter necessitatem’ (‘Don’t multiply entities without necessity’) – except I’ve swapped in ‘assumptions’ for ‘entities’. The classic medieval formulation has its source in a commentary on Duns Scotus by the Irish Franciscan, John Punch (Duns Scotus 1639, book III, dist. 34, q. 1).
quotes the physics educator E. M. Rogers’ description of Ptolemitic astronomy to make this point (Hempel 1966, 231). In contrast to Ptolemy’s “gorgeously complicated system of main circles and sub-circles, with different radii, speeds, tilts, and different amounts and directions of eccentricity” (Rogers 1960, 240), Copernicus’ theory is far simpler.

While it may be unclear exactly what simplicity amounts to, it seems that it bears on the justification of beliefs in non-scientific cases as well. For example, say a detective believes the hypothesis that the murder was committed by the man found at the scene with a knife in his hand, over the hypothesis that the victim’s wife is the murderer and framed the man at the scene. It seems plausible in such a case that the reason the detective’s belief justified is (in part) that the former hypothesis is simpler than the latter. This example, like many others, shows that simplicity bears on justification across the board, and not just in cases of deciding between scientific theories or hypotheses.

Another virtue that is often mentioned is whether a hypothesis or theory is refutable, i.e. whether it provides predictions that can refute or corroborate that hypothesis or theory (Quine and Ullian 1978, 79–80). Popper’s notion of ‘falsification’ is related to this kind of virtue (1959, 17–20).

We can illustrate this virtue, and the corresponding vice of irrefutability with an example of Popper’s. Popper argued that psychoanalysis should be rejected because he claimed it was unfalsifiable. Whenever a psychoanalytic theory made a prediction that was contradicted by the evidence, Popper claimed, the case was reinterpreted so as to make it consistent with the psychoanalytic theory. This resulted, Popper argued, in there never being predictions that could genuinely corroborate or falsify the claims of psychoanalysis (1963, 33–37). Setting aside the question of whether Popper was right about psychoanalysis, it’s clear that he was right that if a theory or hypothesis is irrefutable, then that counts against that theory. It also seems that considerations of refutability are not limited to scientific contexts.

A related and often cited theoretical vice of a hypothesis or part of a theory is if it is *ad hoc*, i.e. if it is only there to support another hypothesis or part of the theory, and has no independent justification (Popper 1963, 60; Quine and Ullian 1978, 78; Bamford 1999). A vivid historical example of an ad hoc scientific hypothesis, which I take from Alan Chalmers, is the Aristotelian response to Galileo’s discovery that there are mountains and craters on the moon, and the moon was not in fact perfectly spherical as Aristotle had claimed. Aristotelians responded to Galileo, as Chalmers describes, by
claiming “that there was an invisible substance on the moon filling the craters and covering the mountains in such a way that the moon’s shape was perfectly spherical” (Chalmers 1999, 76). It is clear that there is no independent support for this hypothesis and that it is only there to preserve the Aristotelian theory. Ad hoc assumptions like this clearly are a theoretical vice, both in the scientific case and across the board.

As I’ve outlined, many of these properties of scientific hypotheses that are described as ‘theoretical virtues’ have a much broader import: they seem to give us a diverse set of properties that in many cases determine whether any belief counts as justified or unjustified. The properties we’ve discussed so far look like they bear on the justification of belief across the board.

There are other so-called theoretical virtues that seem to be much more science-specific than the ones we have covered so far. For example, it is sometimes said that generality or breadth of scope is a virtue of scientific theories (Kuhn 1977, 322; Quine and Ullian 1978, 73–79). A commonly given example of a scientific theory that improves on its predecessors in this regard is Newton’s theory of gravitation. Newton’s theory provides a single explanation of the movements of both terrestrial and heavenly bodies, which is much more general than the explanations provided by earlier hypotheses (Quine and Ullian 1978, 74–75).

Another seemingly science-specific virtue sometimes mentioned is the fruitfulness or fertility of a scientific theory (Kuhn 1977, 322; Newton-Smith 1981, 226–27). A theory is fruitful if it opens up new opportunities for scientific research. As Kuhn puts it, a scientific theory should “disclose new phenomena or previously unnoted relationships among those already known” (1977, 322).

These two virtues, seem particularly tied to the aims of science, do not seem to give us properties that bear on justification across the board. With respect to generality, science is particularly concerned with working out all the different uniformities there are, and how they fit together (Goldman 2001, 60). This aim seems entirely irrelevant in relation to the justification of many non-scientific beliefs, such as beliefs about my current experience. The same seems to be true of fruitfulness. One might prefer seemingly fruitful hypotheses if what one is after is research funding, but it looks irrelevant to the justification of most non-scientific – or, more broadly, non-academic – beliefs.

But setting aside these last two virtues, the first five virtues I discussed –
conservatism, modesty, simplicity, refutability, and non-ad-hoc-ness – provide the basis for another argument in support of the claim that belief is subject to a truth norm. This is because they are a diverse set of properties of beliefs that can bear on the justification of belief. Like with the diverse methods, I'll argue that there must be a unifying reason why all these different properties can justify beliefs. Without such a unifying reason – i.e. if these different properties were radically heterogeneous – then the diversity of these properties would be problematic. Once again, I'll argue that a truth norm, through my TRUTH CONNECTION formula, can be the unifying reason that explains why these different properties can justify beliefs.

2.2. Why something must unify the different properties that bear on the justification of belief

I'm going to argue that there must be something that unifies these different properties that bear on justification. I will again do so by arguing that it would be problematic if these different properties were, in my terms, radically heterogeneous, i.e. if there were no one thing which all these different properties have in common.

We can see this if we consider an alternative view, according to which the different properties that bear on justification do not have anything in common.

This kind of view seems to be put forward by Gilbert Harman, in his discussion of reasoning. Harman describes reasoning as ‘a reasoned change in view’, i.e. a reasonable revision of a belief or addition of a new belief on the basis of beliefs one already holds (1986, 1–4). What makes a revision of one’s beliefs reasonable? Harman lists numerous factors that determine this, including some that we have already mentioned, such as how simple, how conservative and how explanatory the new belief is (1986, chaps. 5–7).

Contrary to what I want to say, however, Harman does not seem to want to give one unifying reason as to why all these factors bear on the reasonableness of belief revision. He definitely does not want to say that all factors that bear on the reasonableness of belief revision have something to do with truth, since he claims that some pragmatic factors determine this as well, such as the interests of the believer and the triviality of the belief (Harman 1986, 55–56).

What advantages does my view – that a truth norm unifies justification – have over a view like Harman’s? There are two main reasons why a view like Harman’s is problematic.
The first main reason is because denying that there is a unifying reason why all these different properties bear on the justification of belief leaves it a total mystery how to weigh these different properties against each other. If all of these properties can bear on the justification of a belief, there must be some way of weighing them against each other in order to work out whether a belief is justified in an all-things-considered sense.

There must be some way to weigh these properties against each other because it seems these properties can push us in opposite directions. For example, a simple belief could be less well supported by the evidence than an alternative, but nevertheless be more justified (as was the case, if Kuhn is correct, with Copernicus’ theory before Kepler).

In different cases, conservatism and refutability can push in opposite directions, i.e. if it’s the case that one can only make one’s belief refutable with a massive revision of one’s beliefs. If Popper is right that psychoanalysis is an unfalsifiable theory, then someone who believes in psychoanalysis would be in such a position.

How is someone to weigh these factors against each other, in order to work out whether a belief is all-things-considered justified? How much evidential support does a non-simple hypothesis need in order for it to be more justified than the corresponding simple hypothesis? How irrefutable does a belief need to be for one to be unjustified in holding it, even given the push of conservatism?

If there is no unifying reason why these different properties all bear on the justification of belief, these questions look almost unanswerable. We can and do weigh these different properties against each other when considering whether a belief is justified, but if there’s no unifying reason why these properties all bear on the justification of belief, it’s a mystery as to how anyone is supposed to be able to do this.

The second main reason why it would be problematic if these properties did not have something in common is that there are some cases in which a belief’s having one of these properties does not seem to justify a belief to any degree.

We can illustrate this with the example of conservatism. Let’s imagine that revising a belief that p requires one to revise many other beliefs. Does that justify one in holding on to the belief that p? The correct answer seems to be ‘only sometimes’. The fact that revising a belief that p would require revising other beliefs does not seem to justify one in continuing to believe that p if the beliefs one would have to revise are the result of, say, a racial bias or reading the tea leaves. In such a case, the fact that one
already believes that p does not seem to justify believing that p to any degree.

The same seems to be the case with simplicity. If one has reason to believe that the facts about a certain subject are complicated, then the fact that a belief about that subject matter is simpler than other hypotheses doesn’t seem to justify that belief to any degree. This is a thought expressed in an aphorism often attributed (perhaps apocryphally) to Einstein: ‘everything should be made as simple as possible, but not simpler’.³

A philosophical example of a case like this can be found in J. L. Austin’s discussion of excuses. Austin claimed that a person’s conduct is such a complicated topic that we should be wary of a simple explanation of the concepts of doing something intentionally, deliberately, and on purpose:

It is worth bearing in mind … the general rule that we must not expect to find simple labels for complicated cases. … [H]owever well-equipped our language, it can never be forearmed against all possible cases that may arise and call for description: fact is richer than diction.  
(Austin 1956, 21)

If Austin is right that we should expect the facts about conduct to be complicated, then the simplicity of a philosophical account of action would not count in its favour.

Of course, someone who holds a view like Harman’s will respond that these properties of beliefs only justify them pro tanto, or prima facie, or all else being equal. But this response just takes us back to our first problem. I then just want an explanation of why conservatism and simplicity do justify beliefs in some cases but not in others – i.e. an account of how one can weigh these different properties against each other.

These two reasons show that it would be problematic if the different properties that bear on justification did not all have something in common. In the next section I’ll argue that belief being subject to a truth norm can explain what all these properties have in common.

2.3. How a truth norm unifies the different properties that bear on the justification of belief

The way a truth norm can explain what unifies the different properties that bear on

³There is no positive evidence that Einstein actually said or wrote this. The first attribution to Einstein of this phrase is found in a New York Times article from 1950 by the composer Roger Sessions (for details, see: http://quoteinvestigator.com/2011/05/13/einstein-simple/).
justification is again through my formula TRUTH CONNECTION:

**TRUTH CONNECTION:** A factor F justifies a belief that p if and only if F counts in favour of p being true.

Before I give this explanation, I should quickly note that I’m not claiming, in this section, that appealing to TRUTH CONNECTION is the only possible way to unify the different properties that bear on justification, but only that a truth norm can unify these properties. I’ll argue afterwards, in §3, that an explanation featuring a truth norm is preferable to other alternative explanations.

Given TRUTH CONNECTION, there is an explanation of why these different properties all bear on the justification of belief. If having a belief with one of these properties is a good strategy for having a true belief in some cases, that explains why that property justifies that belief in those cases. Likewise, if having a belief that lacks one of these properties is a bad strategy for having a true belief in some cases, that also explains why that property justifies that belief in those cases. I’m going to go through the diverse properties that can bear on justification, and suggest cases in which having beliefs lacking each of these properties is a bad strategy for having a true belief in each case.

I am not going to show that a belief’s having these properties counts in favour of its being true in all cases, nor would I want to do so. In fact, I’ll show that beliefs having these properties does not count in favour of those beliefs being true in some cases. But this is grist to my mill, because I’ve already established that, in some cases, a belief’s having these properties does not justify that belief. Therefore, the fact that a belief’s having these properties does not always count in favour of its being true in fact further supports my claim that TRUTH CONNECTION explains why these properties bear on the justification of belief (if, of course, the cases in which a belief’s having one of these properties does not count in favour of its being true just are the cases in which that property doesn’t justify the belief).

Let’s start with conservatism. It seems that if a hypothesis that p is more conservative than another hypothesis q – i.e. if it conflicts with fewer of my other beliefs than does q – this can count in favour of p being true. This is because I don’t just happen to have all the other beliefs I have. I probably have reasons for thinking that they are true or reasons for thinking that they are the result of a reliable mechanism. In such a case, the less conservative hypothesis looks like it’s more likely to be false because it
conflicts with more of my other beliefs.

Davidson gives an example of this kind of choice when thinking about beliefs about the mental states of others, the example of someone who says ‘There’s a hippopotamus in the refrigerator’ (1968, 100–101). This person then tells you that what he calls ‘hippopotamuses’ make great juice, are a great source of vitamin C, and can be bought at the greengrocer. In response to this, there are (at least) two options. I could revise a great many of my initial beliefs about what this person thinks about hippopotamuses; or I could revise my initial belief that this person is referring to hippopotamuses with the word ‘hippopotamus’. Choosing the option that involves revising more beliefs, in this case and many others, looks like a very risky strategy for believing truly. In this way, conservatism can sometimes justify beliefs because it sometimes counts in favour of those beliefs being true.

However, there are cases in which a hypothesis that requires extensive belief revision does not look like a risky strategy for believing truly. Examples of such cases are those in which someone has an extensive worldview that is largely consistent yet not based on good evidence – conspiracy theorists provide good examples of such believers. A conspiracy theorist revising his beliefs about hidden plots both a) looks like he would be following a better strategy for believing truly, but b) looks like he would be engaged in extensive belief revision. And that is exactly as it should be, because in such cases conservatism does not look like it justifies one in believing a hypothesis which fits in with his conspiracist beliefs.

Now let’s consider modesty. It also seems clear that believing a logically stronger hypothesis can be a riskier way of ending up believing the truth than believing a weaker one. Believing that all the GPs in my surgery are male just on the grounds that my GP is – rather than believing something logically weaker – may result in me believing truly. But it’s a very risky strategy for believing truly. Therefore, the fact that a hypothesis is more modest than another can count in favour of its truth, because the more modest hypothesis runs less risk of being false than the less modest one. In this case, TRUTH CONNECTION explains how modesty can sometimes justify beliefs.

This also allows us to explain when modesty does not justify one hypothesis over another. The fact that the proposition that p is logically weaker than the proposition that p & q doesn’t justify believing that p instead of believing that p & q, if one also has reason to believe the proposition p if and only if q. In this case, the modesty of the belief
that p doesn’t make it more likely to be a true belief than the belief that p & q. Therefore, according to TRUTH CONNECTION, modesty does not justify believing that p over believing that p & q in such a case.

Let’s move on to refutability. How does TRUTH CONNECTION explain why refutability bears on justification? It does so because having an irrefutable belief is a very risky strategy for ensuring that one’s belief is true. If I have an irrefutable belief – i.e. one which doesn’t enable me to make any substantive predictions – then it is easy to get away with thinking that I have a true belief when in fact I have a false belief. Say I believe that there is a monster under the bed who becomes invisible whenever I look under there. If I hold such a belief, it’s easy for me with get away with thinking I have a true belief when in fact I have a false belief. In this way, TRUTH CONNECTION also explains why the refutability of a belief can in some cases contribute to its justification.

By modifying this example slightly, I can also show how the ad-hoc-ness of a hypothesis counts against its truth. An ad hoc belief is one that is only held in order to justify holding on to another belief. Say I initially just believe that there is a monster underneath my bed. Then when I look under the bed and see nothing there, I come to believe that the monster must have turned invisible, so I don’t have to revise my initial belief that there was a monster there. Forming a belief in this way, with the sole aim of explaining away recalcitrant evidence, is a bad way of trying to have a true belief. It may in some cases lead to a true belief, but it seems likely in most cases to lead to believing falsely. Therefore, because a belief being ad hoc counts against its truth, this enables TRUTH CONNECTION to explain why the ad-hoc-ness of a given belief detracts from its justification.

This also enables us to say why in some cases seemingly ad hoc beliefs actually are justified. Forming a belief just in order to deal with evidence that seems to disprove an initial belief can look like it a good strategy for believing truly, if there is overwhelming reason to think that the initial belief is true.

Hume’s argument against believing in God on the basis of miracles is such a case. When faced with someone claiming to have witnessed a miracle, it may seem ad hoc to believe, as Hume does, that the witness likely is either lying or has been deceived (1748, 90). (One can imagine the witness saying “Why, Hume, did you believe me when I told you what the time was earlier, but you don’t believe me now?”). But Hume’s belief is not in fact ad hoc, because miracles are events which we have overwhelming reason to think
do not occur. So the fact that Hume’s belief is formed in order to deal with recalcitrant
evidence in this case counts in favour of its truth, given the overwhelming reason there is
in thinking that miracles do not occur. So Truth Connection can explain when the
seeming ad-hoc-ness of a hypothesis does not undermine its justification as well.

Now let’s consider simplicity. I’ll have to spend a bit more time on this because,
as I noted earlier, simplicity is less well defined than the other virtues.

Because of this, all I can do here is suggest that in some cases the simplicity of a
hypothesis does count in favour of its truth. This is because in some cases appeals to
simplicity just are appeals to conservatism or modesty. I’ll also suggest that in certain
other appeals to simplicity – appeals to simplicity in deciding between scientific theories
– it looks like a virtue that is particularly linked to the aims of science, and not something
that bears on justification across the board.

If we go by something like Ockham’s Razor – ‘avoid postulating assumptions
without necessity’ – we can say a hypothesis is simpler than another if involves fewer
assumptions.

With this in mind, a simple hypothesis could in some cases just be a conservative
hypothesis or a modest hypothesis. A hypothesis that is more conservative than another
could count as simpler, because it involves fewer assumptions that conflict with our previous
assumptions. A hypothesis that is more modest than another could also count as simpler –
because it involves fewer assumptions full stop. And we have already shown how each of
these two features can in some cases count in favour of a belief’s truth.

Is this what people mean when they speak of favouring a hypothesis over
another on grounds of simplicity? There are definitely cases in which what is called a
‘simple’ hypothesis just is a conservative hypothesis. Davidson, in the example we
discussed above, describes the hypothesis that the person does not mean what I mean by
the word ‘hippopotamus’ as “the simpler hypothesis by far” (1968, 101). And that was a
case of preferring a hypothesis on the grounds of conservatism.

On the other hand, ‘simplicity’ sometimes just refers to modesty, i.e. a hypothesis
is simpler than another if it involves fewer assumptions full stop. This understanding of
simplicity is implicit in the argument that a simpler hypothesis has an inherently higher
probability than a more complicated one. Elliott Sober summarises this kind of
argument, which he attributes to Russell and Quine, as the claim “that removing an
existential claim from a theoretical system has the effect of raising the probability of what
remains ... because a conjunction must have a lower probability than either conjunct, provided that the conjuncts are mutually independent” (Sober 1981, 145). If this is supposed to show how *simple* hypotheses are preferable, it can only do so on the assumption that simplicity is modesty.

So ‘simplicity’ sometimes means conservatism and sometimes means modesty, and we’ve already shown how TRUTH CONNECTION explains why both properties can justify beliefs in some cases.

On the other hand, there are also cases in which ‘simplicity’ is a virtue of scientific theories in a science-specific way.

For example, Newton’s theory of gravitation is often claimed to have the virtue of being simpler than previous theories because it gave a single explanation of the movements of any kind of body, whether terrestrial or heavenly (see e.g. Hempel 1966, 234–35; Quine and Ullian 1978, 70). Appeals to the simplicity of Newton’s theory look very different from the appeals to simplicity I’ve just discussed. It’s much harder, if not impossible, to see how the simplicity of Newton’s theory counts in favour of its truth.

This is not, however, a problem for me, given the sense of justification I’m concerned with. I’m concerned with a very general sense of justification that can apply across scientific and non-scientific contexts. The simplicity of Newton’s theory seems to be a virtue that is particularly tied to the aims of science, in that science aims at discovering *regularities*, or providing *unifying* explanations of disparate phenomena (cf. M. Friedman 1974). These kinds of science-specific aims do not look relevant to the justification of belief across the board.

The three different ways of appealing to simplicity I have just sketched do not say everything about how simplicity bears on the justification of either everyday or scientific beliefs. But I want to draw a tentative conclusion on the basis of them.

On the one hand, the first two ways of appealing to simplicity – when simplicity just means either conservatism or modesty – are ways in which it does bear on the justification of beliefs across the board. They are also cases in which TRUTH CONNECTION can explain *why* they bear on justification, because conservatism and modesty do in some cases count in favour of truth.

On the other hand, the last way of appealing to simplicity – when being ‘simple’ means being general or unifying disparate phenomena – does not seem to bear on the justification of beliefs across the board. My tentative conclusion is that that insofar as
simplicity bears on the justification across the board, it is because simplicity counts in favour of truth.

This concludes the defence of my claim that each of these different properties can bear on the justification of a belief because they can all bear on whether or not that belief is true.

Furthermore, it should also be clear how TRUTH CONNECTION explains why it’s possible to weigh each of these different properties against each other. If all of these properties justify beliefs to the extent that they count in favour of them being true, then it doesn’t look mysterious how we can, say, play considerations of conservatism off against considerations of refutability. In fact, I’ve been doing something like that in outlining some of the cases in which the fact that a belief has one of these properties does not justify that belief.

To say that it becomes unmysterious why we can weigh these different properties against each other is not to say that it becomes easy or trivial. It’s often hard to figure out when a belief is sufficiently irrefutable that it becomes unjustified, even given the fact that it is consistent with many other of one’s beliefs. But the explanation I’ve given does make sense of how it is possible to weigh these properties against each other. Since we can and do weigh these different properties against each other when considering whether a belief is justified, we’d better be able to explain how is possible to do so.

In summary, thinking of belief as subject to a truth norm can explain – through my formula TRUTH CONNECTION – what unites the different properties that bear on justification. Therefore, if belief is subject to a truth norm, then these properties are not problematically heterogeneous, as they were on Harman’s view.

3. Alternative explanations of what unifies justification

I’ll now end this chapter by discussing some alternative explanations of what unifies justification. In particular, I’ll discuss whether a) a knowledge norm, or b) an evidence norm can unify justification. I’ll argue in both cases that there is reason to prefer an explanation featuring a truth norm.

3.1. A knowledge norm

Consider the following knowledge norm:
KNOWLEDGE: One ought to believe that p if and only if one knows that p.

Some claim that a knowledge norm like this – rather than a truth norm – is the fundamental norm on belief (Sutton 2005; 2007; Bird 2007; Littlejohn 2013). This is seen by its proponents as part of the project of ‘knowledge first’ epistemology pioneered by Timothy Williamson (2000).

I’m not going discuss the reasons why these philosophers prefer a knowledge norm over a truth norm. These reasons often include claims that a truth norm cannot explain justification, so that we have to appeal to a knowledge norm instead (see e.g. Bird 2007, 93–95; Littlejohn 2013). I’m not going to critique these reasons here (though for criticism, see Whiting 2013a). This is because in Part Two of this thesis I’m going to outline the way in way a truth norm can succeed in explaining justification. If this account is correct, then these philosophers must be wrong.

In this chapter, I am just going to discuss which out of a knowledge norm and a truth norm is better served to explain what unifies justification. One might think that a knowledge norm could also explain what unifies justification, i.e. that it can play the unifying explanatory role which I’ve suggested a truth norm plays. According to such a picture, what would unify all the different methods and properties that can justify a belief that p would be that all of them in some way support one’s knowledge that p.

I think such an approach to what unifies justification is problematic in a way that my truth norm approach is not. This is because a knowledge norm faces the same problem as the trivially unifying explanations we discussed above.

One example of such a trivially unifying explanation was the claim that what unifies different methods and properties which justify a belief that p is that they provide reason to believe that p. We saw above that this claim was completely unexplanatory, because it pointed to a feature that justified beliefs trivially share. Saying that a method provides one with reason to believe that p didn’t seem to add anything to saying that the method justifies one in believing that p. Because of this, this claim didn’t look like it could provide us with an informative story about how all the different methods can interact and how all the different properties can be weighed against each other.

A knowledge norm looks like it will be equally uninformative in this context. Let’s grant to the proponents of a knowledge norm that all factors that support knowledge that p are factors that justify one’s belief that p. Even granting this, it looks likely to fail in a similar way to the attempted explanation in terms of reasons for belief.
For example, it doesn’t seem like it will be able to explain, or give an informative story about, how different methods can interact in supporting and defeating justification.

We can see this by returning to the example of Maryam Mirzakhani. When Maryam Mirzakhani corrects my mental arithmetic, it seems that the justification for my belief that the hypotenuse was 10cm – which was formed by mental arithmetic – was defeated by a completely different method, i.e. her testimony. Merely pointing to the fact that both mental arithmetic and testimony can support and undermine knowledge doesn’t provide the explanation we want. It doesn’t explain, that is, why a belief justified by one method can have its justification undermined by a completely different method. We are still left wondering why both mental arithmetic and testimony can undermine or support knowledge in the same way, and how these methods are supposed to interact (just as we were earlier left wondering why these distinct methods could equally provide reasons for belief, and how they were supposed to interact). This question is left no less mysterious by a knowledge norm approach, as it was by a reasons-for-belief approach.

A truth norm, as I have shown, can explain what unifies justification. A truth norm can provide us with a genuinely informative and explanatory story about how different methods and properties that justify beliefs interact. In this way, the fact that it can explain the unity of justification is a point in favour of a ‘truth first’ approach to epistemology rather than a knowledge first approach.

The defender of a knowledge norm could object at this point by appealing to a particular feature of knowledge, which can provide an informative story about how different justifying methods and properties interact. For example, some proponents of knowledge-first epistemology, most notably Williamson (2000, 149), still claim that knowing that p entails that one has a belief that p that is ‘safe’, i.e. a belief that one would still have in nearby possible worlds in which p is still true.

The defender of a knowledge norm could claim that safety can provide us with a feature that provides an informative story about what unifies the different methods and properties that bear on justification. Such an account would claim that something only bears on whether a belief is justified insofar as it counts in favour of, or against, that belief being safe. This may seem to be more informative than merely appealing to a knowledge norm in the way I discussed above. Since beliefs ought to be safe because they ought to be knowledge, according to such an account, this is how a knowledge norm can get us an informative story about what unifies justification.
I agree with the part of this objection which claims that appealing to safety could be more informative than merely appealing to a knowledge norm in the way I discussed above. However, according to such an account it would be a safety requirement — rather than a knowledge norm — which is doing all the explanatory work. Furthermore, it doesn’t seem that safety would have to be explained by a knowledge norm. Indeed, it seems that a safety requirement could equally be explained by a truth norm. Therefore, it doesn’t seem that appealing to a safety requirement can show that a knowledge norm — rather than a truth norm — can do the requisite explanatory work.

3.2. An evidence norm

Consider the following evidence norm:

EVIDENCE: One ought to believe that p if and only if p is supported by one’s evidence.

Some philosophers claim that a norm like this is the fundamental norm on belief (see e.g. Feldman 1988, 254; 2000, 678–79). This forms part of a more general claim, known as ‘evidentialism’, that the concept of a belief’s being supported by the evidence is the fundamental epistemic concept (Conee and Feldman 1985).

One might think that an evidence norm could also explain what unifies justification in the way I’ve suggested a truth norm can. According to such a picture, what would unify all the different methods and properties that can justify a belief that p would be that all of them in some way provide evidence that supports p.

I don’t think that such a picture is plausible. To show this we need to distinguish two broad ways in which we could understand what counts as evidence on this picture. Firstly, it could be what’s referred to by the common-sense concept of evidence. Secondly, evidence could be understood probabilistically. Both understandings of evidence have problems. With the former, the common-sense concept of evidence is far too specific to explain what unifies justification. With the latter, understanding evidence

---

4 I won’t attempt to do this because, as I said in Chapter 1, I am concentrating on whether a truth norm can explain justification, and setting aside the question of whether it can explain knowledge. Though for a defence of the claim that a truth norm can explain knowledge as well (and something like a safety requirement), see Wedgwood 2002, secs. 5–6.
probabilistically has a leftover issue about how much the evidence needs to support p for one to be justified in believing that p.

3.2.1. The common-sense concept of evidence

By the common-sense concept of evidence, I mean the concept as used in common parlance when people speak of having (or not having evidence) for what they believe. If we understand evidence in this way, then it seems far too specific to explain what unifies all the different methods and properties that justify beliefs. This is because there are some kinds of beliefs which it sounds wrong to say we have evidence for, but it doesn’t sound wrong to call them justified. I’ll give three examples.

Firstly, in most cases it sounds strange to say that we have evidence for our beliefs about mathematical matters, even when they are justified. It does look possible in some cases to have ‘evidence’ about mathematical matters. In my example of Maryam Mirzakhani correcting my mental arithmetic, it looks like her testimony gave me evidence that the hypotenuse was 9cm. But if I come to a mathematical belief on the basis of mental arithmetic or by carrying out a proof, it sounds very strange to say I have ‘evidence’ for my conclusion. But it doesn’t sound strange say that a mathematical belief based on mental arithmetic or proof is justified.

Secondly, it (typically) sounds strange to say that we have evidence for those of our beliefs which form part of our self-knowledge, e.g. beliefs about our own current mental states, the positions of parts of our bodies, and our own intentional actions. To illustrate, it sounds extremely strange to say that I have evidence for my belief that I want a coffee, my belief that I am sitting down, or my belief that I am writing a thesis. However, it does not – to me at least – sound strange to say that such beliefs are justified.

There are, admittedly, some who would deny that the beliefs which form part of our self-knowledge are justified. Sometimes this is claimed because it is implausible that such beliefs are based on evidence. For examples of philosophers who make this claim about our beliefs about our intentional actions, see (Hampshire 1975, 53; Roessler 2013). Examples of philosophers who make this claim about our beliefs about our own mental states include some expressivists about self-knowledge (see e.g. Bar-On 2004, chap. 9), and also those who claim that we are ‘entitled’, in the sense introduce by Burge (1996), to beliefs about our own current mental states, although they are not justified (see e.g.
Peacocke 2009).

In the context of this chapter, I do not want to deny that these are possible (and even plausible) views. But I do think it would be best if the evidentialist didn’t have to appeal to a view like one of these merely in order to defend evidentialism. That would seem ad hoc.

Thirdly, it sounds strange to say that we have evidence for our beliefs based on direct perceptual experience. This is a point Austin makes when he asks what kind of things are normally said to count as evidence that an animal in a particular enclosure is a pig. The natural kind of things we would understand as evidence that there’s a pig would be things like a half-eaten vat of swill, or snout- or trotter-like marks in the mud. But, as Austin puts it:

If the animal then emerges and stands there plainly in view, there is no longer any question of collecting evidence; its coming into view doesn't provide me with more evidence that it’s a pig, I can now just see that it is, the question is settled. (Austin 1964, 115)

I agree with Austin that it is strange to think of direct perceptual experience as just providing us with more evidence. But again I don’t think it is strange to think of a belief based on direct perceptual experience as justified.

Therefore, if evidence is understood as what is referred to in common parlance when people speak of having or not having evidence for what they believe, then these three kinds of belief show that it is then far too specific to unify justification.

This is of course not to say that we can’t come up with an account of evidence on which we do, despite appearances, have evidence for all of these three kinds of beliefs. The probabilistic account of evidence, which I’ll discuss next, is a very popular way to do this. All I’ve shown so far is that if an evidence norm is going to unify justification, an additional theory of evidence of this kind is required; it is not part of, and would be somewhat in conflict with, the common-sense concept of evidence.

3.2.2. A probabilistic conception of evidence

A probabilistic conception of evidence claims that how strongly a body of evidence E supports p is how likely p is given E. Can adopting a probabilistic conception of evidence result in an evidence norm that can succeed in unifying justification? According to such a picture, what would unify all the different methods and properties that can justify a belief that p would be that all of them raise the probability of p.
The reference to probabilities in this picture can be understood either subjectively or objectively. Those who understand them subjectively understand these probabilities just as the actual degrees of confidence of subjects, otherwise known as ‘credences’ (see e.g. Jeffrey 1992, esp. 196-98). Those who understand them objectively think that something only counts as evidence for p if it in some sense genuinely raises the probability that p. One way to do this is to claim that E only provides evidence for p if one knows that E raises the probability of p (see e.g. Maher 1996; Williamson 1997). I’m going to remain neutral on this question because it is not relevant to the problem I want to raise about whether a probabilistic evidence norm can unify justification.

The problem I want to raise for a probabilistic evidence norm is different from the problems I’ve raised for the other alternative norms on belief I’ve discussed so far. I do not want to argue that such a norm would fail to unify justification in the same way as these alternative norms, because I admit that the probabilistic conception of evidence has better prospects of unifying justification than the common-sense concept of evidence. Whatever the reasons are for my mathematical beliefs, my beliefs about my own mental states, and my beliefs based on direct perceptual experience, it does seem plausible that they raise the probabilities of those beliefs. I myself will appeal to this broader probabilistic conception of evidence later, in Chapter 6, as I think it has to be part of an explanation of justification in terms of a truth norm.

To see the problem with a probabilistic evidence norm, we will have to take a detour away from the question of what unifies justification. A probabilistic evidence norm has a problem in relation to the question of how much one’s evidence needs to support p in order for one to be justified in believing that p. The evidentialist has to admit that in order to have a justified belief that p, it is not enough just to say that one’s evidence supports p somewhat. Rather, in order to have a justified belief, one’s evidence needs to support p above a certain threshold that is enough or sufficient for justification.

If a probabilistic evidence norm really is the fundamental norm on belief that explains justification, it’s going to have to provide an answer to this question as well, i.e. it’s going to have to tell us what makes for evidence that is sufficient for justification.

However, I don’t think a probabilistic evidence norm can provide a good answer to this question. I’ll show this by outlining three different possible answers, all of which look inadequate.

Firstly, the simplest way we could provide an answer to how much evidence is
sufficient for justification would be just to claim that one has enough evidence if \( p \) is the most likely possibility given one’s evidence, as expressed in the following norm:

**MORE THAN 50-50:** One ought to believe that \( p \) if and only if one’s evidence makes \( p \) more likely than not-\( p \).

However, such an evidence norm simply looks false. It’s easy to find examples in which \( p \) is the most likely possibility given one’s evidence, but in which one would not be justified in believing that \( p \). To take an example from Michael Pace (2011, 246), say there’s an unfair coin which lands heads 55% of the time. That the coin will land heads on its next toss is the most likely possibility given my evidence, but it is extremely implausible that I would therefore be justified in believing that it will land heads on my next toss. Therefore, this simple account of sufficient evidence won’t work.

A second way of saying how much evidence is sufficient for justification would be just to insert the requirement that one must have sufficient evidence into the content of the evidence norm. The following two evidence norms are examples of such an approach:

**SUFFICIENT EVIDENCE:** One ought to believe that \( p \) if and only if one has sufficient evidence that \( p \).

**EVIDENCE-THRESHOLD:** One ought to believe that \( p \) if and only if, given one’s evidence, the probability of \( p \) is above a particular probability threshold, \( T \).

I am imagining that such an approach would not say too much about what particular probability threshold \( T \) is, or about how much one’s evidence needs to support \( p \) for it to be ‘sufficient’. I assume that we all have an intuitive understanding of what this threshold is.

The problem with this approach is that it is completely unexplanatory. Neither of these two norms give any explanation as to why the threshold for sufficient evidence is where it is, or why it is higher than what more than 50-50 requires. Neither of these two norms can provide any answer to these questions, other than merely saying that the threshold is where it is *because that’s the sufficient threshold for justification*. Given this, if someone were to put forward one of these two norms, one would suspect that it’s the concept of *justification*, rather than the concept of *evidence*, that is doing the heavy lifting. Therefore, merely inserting the concept of sufficient evidence into an evidence norm is
also inadequate.

A third way of explaining how much evidence is sufficient for justification would be to appeal to the relation between all-out beliefs and credences. Consider the following evidential norm which applies to credences rather than all-out beliefs:

EVIDENCE-CREDENCE: One ought to proportion one’s credences to the evidence.

This norm means that if, for example, the evidence one has makes p 0.75 likely, then one ought to have credence 0.75 in p.

Such an evidence norm cannot, on its own, explain how much evidence for p is sufficient for one to be justified in having an all-out belief that p, because it says nothing about all-out belief. However, with an additional psychological assumption, it seems that EVIDENCE (CREDENCE) can answer this question.

The requisite psychological assumption is the so-called ‘threshold view’ of all-out belief. This is a view that all-out beliefs are reducible in a particular way to credences. We need to assume that what it is to have an all-out belief that p just is to have a credence in p above a certain threshold (see e.g. Sturgeon 2008, 141–42).

Given the threshold view of belief, EVIDENCE-CREDENCE can give a possible explanation of how much one’s evidence needs to support p in order for one to have a justified belief that p. On such a picture, evidence that’s sufficient in believing that p just is evidence that means that p is above a particular threshold of likelihood. Specifically, one’s evidence needs to make it likely enough that the credence one ought to have — according to EVIDENCE-CREDENCE — is above the threshold for all-out belief — according to the threshold view of belief.

This is, in essence, a version of what Richard Foley has called the ‘Lockean thesis’:

[I]t is rational for you to believe a proposition just in case it is rational for you to have a sufficiently high degree of confidence in it, sufficiently high to make your attitude towards it one of belief. (Foley 1993, 140)

The difference between Foley’s phrasing and mine is that he states it in terms of ‘rationality’, whereas I’ve stated it in terms of justification, though the fundamental point is essentially the same.

This approach does not have the flaws of the previous two explanations. Firstly, the threshold view of belief typically claims that the threshold for all-out belief is a fair
bit above 0.5 credence in \( p \), and so it does not have the same problem as more than 50-50. Secondly, the threshold for sufficient evidence is not left unexplained on this picture; it’s explained by a psychological fact about what the threshold for full belief is.

I don’t think, however, that this is a good explanation of when one has enough evidence for all-out belief. The chief reason is that the threshold view of belief is an implausible assumption which we would be better off not appealing to.

The threshold view of belief is implausible because we are much more familiar with all-out belief. It is far less clear that credences are mental states we actually have. This makes credences ill-suited to the job of providing a reduction of all-out belief.

That we are more familiar with all-out belief is manifested by the fact that it is the doxastic attitude that primarily figures in our commonsense-psychological explanations. We tend to say ‘He went to the post office because he thought [believed] that he could get stamps there’. In this way, all-out belief definitely seems to be an attitude that our commonsense psychology commits us to.

Numerically valued credences, on the other hand, don’t primarily figure in our commonsense-psychological explanations, but rather look like a model used to account for particular kinds of behaviour such as betting.

This means that the kind of reductive thesis that is embodied in the threshold view of belief is a very strange one. It reduces a mental state we are very familiar with – i.e. all-out belief – to a mental state we are much less familiar with – credence.

Furthermore, the psychological limitations of believers counts against belief being reducible to credence. I’ll outline two ways in which these limitations count against credence being fundamental, both of which I take from Richard Holton (2014, 13–25).

Firstly, and more concretely, it is well documented that we systematically make certain kinds of errors in probabilistic reasoning. For example, when thinking about how likely \( p \) is given some evidence, we often ignore or don’t fully appreciate base-rates (Gigerenzer 2008, chap. 9). We also make the so called ‘conjunction fallacy’ in which we think that the probability of \( p \& q \) can be higher than the probability of \( p \) (Kahneman 2011, chap. 15). If belief were reducible to credence, an essentially probabilistic state, we wouldn’t expect people to be quite so bad at probabilistic reasoning.\(^5\)

\(^5\) In his fuller argument, Holton argues that the base-rate fallacy provides reason to think that we only ever work with beliefs with probabilistic contents rather than with credences. This is because we are less prone
Secondly, and more abstractly, we are creatures who have limited time and cognitive capacities for information processing. It seems that having credences would involve a heavy psychological burden, as Holton outlines in the following passage:

Unless their powers of memory and reasoning are very great, those who employ credences risk being overwhelmed by the huge mass of uncertainty that the approach generates. First they will have to store very much more information: rather than just discarding the propositions that aren’t believed and focussing on those that are, they will have to keep track of all of them and their associated credences. Second, they will have to be able to deploy the complicated methods needed for probabilistic reasoning. The problem will be all the worse if, in Bayesian fashion, they update their credences by conditionalization. For then, in addition to keeping track of the probabilities of the various events, they would also need to keep track of the conditional probabilities of events upon each other: the probability of p given q. (Holton 2014, 14)

Having all-out beliefs, on the other hand, imposes a much more manageable psychological burden on us. In thinking about whether I believe something is the case, this seems to be a fairly simple question with only a very limited set of options, i.e. I either believe it, disbelieve it, suspend judgement, or don’t have any attitude at all. Therefore, given the difference in the psychological burdens that all-out beliefs and credences involve, the latter look ill-suited to the job of providing a reduction of the former.

In this way, the evidentialist has to rely on an implausible psychological assumption – the threshold view of belief – in order for EVIDENCE-CREDENCE to explain what makes for sufficient evidence for justification.

In summary, I have considered three ways in which the proponent of a probabilistic evidence norm can attempt to explain what makes sufficient evidence for justification, and I have shown them all to be inadequate. This should make us sceptical, I suggest, about whether the fundamental epistemic norm is an evidence norm.

to ignore base rates when information is presented in terms of frequencies than in terms of conditional probabilities. This, Holton concludes, gives us reason to doubt that we ever have attitudes that can correctly be thought of as credences (2014, 21–24). If Holton is correct about this, then it lends weight to my argument. However, my argument against EVIDENCE-CREDENCE doesn’t require me to argue for the claim that we never have credences, only the weaker claim that belief is not reducible to credences (which we might nevertheless have).
4. Conclusion

Now I’ve finished my detour into what makes for sufficient evidence, I can summarise what this chapter has shown. I’ve shown that the major alternative norms on belief – a knowledge norm and an evidence norm – have problems a truth norm does not. Therefore, I think we should stick to my initial conclusion that a truth norm provides the best explanation of what unifies justification.

However, my full defence of this claim will be incomplete until I can show that a truth norm can explain what the last alternative norm – a probabilistic evidence norm – failed to, i.e. how much justification one needs to have in order to have a justified belief that p.

I think a truth norm does look better suited than a probabilistic evidence norm to giving such an explanation. This is chiefly because a truth norm is essentially a norm about one ought to all-out believe. Therefore, a truth norm looks much more likely to give us a story about how there is some threshold amount of justification one needs for a justified all-out belief.

However, I don’t think this explanation will be simple; rather, I think it requires detail and some amendments to the basic truth norm story. But I do think it can be done, and I’ll go into the details in Chapter 6, when I discuss suspension of judgement.
PART TWO

HOW A TRUTH NORM EXPLAINS JUSTIFICATION

Part One of this thesis showed that a truth norm could have significant explanatory power within epistemology. Therefore, if there is a defensible version of the truth norm account of justification, that would be a very good thing.

Part Two of this thesis argues that there is indeed a defensible version of the truth norm account of justification. It does so by discussing a number of problems that the truth norm account of justification has been alleged to suffer from. In Chapter 4, I discuss the claim that a truth norm cannot guide believers. In Chapter 5, I discuss how to best formulate a truth norm that does not have the problematic implication that one ought to believe all the truths. In Chapter 6, I discuss the problem created by the fact that a truth norm never prescribes suspending judgement. In Chapter 7, I discuss the claim that a truth norm has the problematic entailment that one ought to make ‘sacrifices for the greater epistemic good’, i.e. that one ought to have some beliefs that look prima facie unjustified, but result in one having a large number of other true beliefs.

None of these problems, I suggest provide us with reason to reject the truth norm account of justification. However, in providing an answer to the last three problems, I’ll need to make modifications and amendments to the basic truth norm picture. It’s only with these modifications and amendments that a truth norm can both deal with these problems and still succeed in explaining justification. It is through answering the various problems a truth norm is alleged to face, therefore, that we can see the contours of a defensible version of the truth norm account of justification.
In this chapter, I'll discuss what has been called ‘the no guidance argument’. This argument, given by Katrin Glüer and Åsa Wikforss, has the conclusion that we should reject a truth norm because such a norm cannot guide belief formation. I'll argue that this argument fails. Drawing on a response given by Asbjørn Steglich-Petersen, I'll argue that Glüer and Wikforss rely on an overly restrictive conception of what it is for a norm to provide guidance. We should, I argue, prefer a more inclusive conception of guidance Steglich-Petersen puts forward, which allows a truth norm to guide. In the end, however, although I do think it is necessary to respond to Glüer and Wikforss’s argument, I don’t think that guidance is a central issue. This is because a truth norm guiding directly is not central to the truth norm account of justification.

1. The no guidance argument

Glüer and Wikforss argue that we should reject a truth norm, because it is not the kind of norm that can guide believers. We can illustrate this with one of the simpler norms I’ve been discussing, OUGHT:

*OUGHT*: One ought to believe that p if and only if p is true.

OUGHT is a norm which tells one to do something – i.e. believe that p – given a certain antecedent condition – i.e. if (and only if) p is true. It’s therefore a norm of the form ‘Do X when in C’. Glüer and Wikforss argue that if we think about how norms of this form guide subjects, we recognise that OUGHT cannot guide in this way. In fact, since all truth norms are norms of the form ‘Do X when in C’, if Glüer and Wikforss’s argument works against OUGHT, it works against any formulation of a truth norm. Therefore, although I’ll end up, in Chapter 5, defending an alternative truth norm to OUGHT, I still need to answer Glüer and Wikforss’s argument.

For a norm to provide guidance, in the sense Glüer and Wikforss are interested in, is for it to provide a subject with motivating reasons. How does a norm of the form ‘Do X when in C’ provide me with motivating reasons? Glüer and Wikforss claim that this involves me forming a belief about whether the norm’s antecedent condition is met,
i.e. a belief about whether I am in circumstances C (Glüer and Wikforss 2009, 44; Glüer and Wikforss 2010, 758). If I believe that I am indeed in circumstances C, then that belief, together with the norm ‘Do X when C’, provides me with motivating reason to do X.

Glüer and Wikforss illustrate this with the example of the stock market trader’s norms **BUY LOW** and **SELL HIGH**, which Boghossian uses to explain how a truth norm is supposed to explain subjective epistemic norms. If I’m a trader on the stock market, and I’m deliberating about whether to sell some stocks, in order for the norms **BUY LOW** and **SELL HIGH** to provide me with guidance, I need to form a belief about whether the market is at a high. If I believe that the market is indeed at a high, that belief, together with the norm **SELL HIGH**, provides me with motivating reason to sell (Glüer and Wikforss 2010, 758–59).

We should note that being guided by the norm **SELL HIGH** doesn’t necessarily mean that I’ll have normative reason to sell. My belief that the market is at a high may be false, which may mean I don’t have normative reason to sell. But we’re concerned with how norms provide guidance or motivation, and my belief about the market doesn’t have to be true in order for this norm to guide or motivate my behaviour. As Glüer and Wikforss put it, “[t]his belief may of course be false, but this does not prevent the rule from influencing, or motivating, my behaviour. … Guidance does not necessarily amount to correct performance” (2010, 758).

Glüer and Wikforss argue that if we apply this picture of guidance to **OUGHT**, it does not look like it can provide guidance in the same way as other norms. If the preceding picture of guidance is correct, what would be involved in being guided by **OUGHT** in deliberating about whether to believe that p? Since **OUGHT** is a norm of the form ‘Do X when in C’, being guided by **OUGHT** would involve forming a belief about whether the norm’s antecedent condition is met. But in the case of **OUGHT**, and indeed in the case of any truth norm, a belief about whether the norm’s antecedent condition is met just is a belief as to whether p. Therefore, in order to be guided by **OUGHT**, I need to form a belief as to whether p. Glüer and Wikforss claim that this “makes it intuitively very strange to think of [OUGHT] as guiding belief formation” (2013a, 83), for two related reasons.

Firstly, Glüer and Wikforss claim that this means that the so-called guidance that **OUGHT** supplies inevitably comes ‘too late’. This is because in order to be guided by
OUGHT, I have to *already* have a belief as to whether p. This makes it strange to think of OUGHT providing guidance because “[t]he very question we wanted guidance on – whether to believe that p – needs to have been answered before any such guidance can be received” (2013a, 84).

Secondly, Glüer and Wikforss claim that this means that OUGHT can never ‘make a difference’ as to what I believe. Because following it has to involve forming a belief as to whether p before OUGHT can provide guidance, “[i]t never ‘tells me’, or ‘provides me with a reason’ to believe other than I would come to believe anyway” (Glüer and Wikforss 2013a, 84).

These two reasons together, Glüer and Wikforss claim, make it strange to think of OUGHT, or indeed any truth norm, as providing guidance. And if it’s not the case that a truth norm can provide guidance, then we should reject the claim that belief is subject to a truth norm.

2. Steglich-Petersen’s response

As I outlined above, I’m going to appeal to a response to Glüer and Wikforss given Steglich-Petersen. Steglich-Petersen claims that the no guidance argument relies on an overly restrictive conception of guidance, and that embracing a more inclusive conception allows us to maintain that truth norms can guide belief formation.

Steglich-Petersen’s response begins by noting that there is a very intuitive sense in which OUGHT does provide guidance. If I’m deliberating about whether to believe that p, OUGHT provides guidance by telling me that I ought to do so if and only if p is true. This intuitive sense in which OUGHT provides guidance can be seen, Steglich-Petersen claims, by distinguishing it from different possible norms on belief which would offer different guidance, such as the following norm:

**PLEASANT**: One ought to believe that p if and only if believing that p would be pleasant. (Steglich-Petersen 2010, 754; 2013a, 280–81)

Intuitively, PLEASANT provides very different guidance to OUGHT. If I’m deliberating about whether to believe that p, pleasant tells me to do something very different from what OUGHT tells me to do. The question is whether we can maintain that OUGHT provides guidance in this intuitive sense in response to the no guidance argument, and Steglich-Petersen argues that we can.
How can we maintain, in response to the no guidance argument, that OUGHT guides in this intuitive sense? Steglich-Petersen claims that we can do so, because we can a) accept Glüer and Wikforss’s claim that being guided by OUGHT involves forming a belief as to whether the norm’s antecedent condition is met, i.e. a belief as to whether p, but b) deny their claim that this means that OUGHT can only provide guidance after I’ve formed a belief that p. Instead, OUGHT can provide guidance by governing the way in which I form a belief about whether the norm’s antecedent condition is met – i.e. by governing the way in which I form a belief about whether p.

We put this by saying that although being guided by OUGHT involves, as a necessary condition, forming a belief as to whether p, this is not sufficient to count as being guided by OUGHT. Whether one counts as guided by OUGHT also depends on how one forms a belief as to whether p. In particular, Steglich-Petersen claims, one only counts as guided by OUGHT if one forms “a belief as to whether p, as part of an effort to determine whether p is true” (2013a, 281).

If, on the other hand, one does not form a belief as to whether p as part of an effort to determine whether p is true, then one does not count as guided by OUGHT. For example, if one forms a belief as to whether p in a way that is unconcerned or uncaring for whether p is true, then one is not guided by OUGHT. Perhaps, in some cases, one could count as guided by PLEASANT instead, if one formed a belief as to whether p as part of an effort to determine whether believing that p would be pleasant. (Admittedly, as Steglich-Petersen admits, it would be hard to imagine consciously being guided by PLEASANT in forming a belief. However, it’s sufficient to make the point if one were to be guided by PLEASANT in a subconscious or indirect way (Steglich-Petersen 2013a, 282, fn.3)).

Therefore, on Steglich-Petersen’s picture, the guidance that OUGHT provides is not a guidance that happens after one has formed a belief as to whether the antecedent condition has been met, it’s a guidance that is involved in the way in which one forms a belief as to whether the antecedent condition is met. One only counts as guided by OUGHT if one forms a belief about whether p is true as part of an effort to determine whether p is true.

Is what Steglich-Petersen outlines here a genuine way in which a truth norm like OUGHT can guide belief formation? Steglich-Petersen argues it is. He does so by distinguishing between two different conceptions of guidance. Firstly, there is a
narrower, more restrictive sense of guidance, which Glüer and Wikforss are assuming, and according to which what Steglich-Petersen has outlined does not count as guidance. Secondly, there is a wider, more inclusive conception of guidance, according to which what Steglich-Petersen has outlined does count as guidance, and which he claims we should adopt. Steglich-Petersen defines these different conceptions of guidance with the following competing conditions on guidance:

**NARROW GUIDANCE**: A norm N of the form ‘Do X when in C’ can guide S’s behaviour with respect to X only if S following N can make a difference to S X-ing after S has formed a belief as to whether C obtains. (Steglich-Petersen 2013a, 282)

**WIDE GUIDANCE**: A norm N of the form ‘Do X when in C’ can guide S’s behaviour with respect to X only if S following N can make a difference to S X-ing. (Steglich-Petersen 2013a, 281)

The question of whether we should adopt NARROW GUIDANCE or WIDE GUIDANCE is a question about how many ‘steps’ there need to be involved in a norm providing guidance. NARROW GUIDANCE requires there to be two steps involved in a norm providing guidance, because it holds that a norm can only provide guidance if it makes a difference to a subject’s behaviour only after that subject has formed a belief as to whether the norm’s antecedent condition has been met. In this way, NARROW GUIDANCE holds that forming a belief as to whether a norm’s antecedent condition is met and the norm’s making a difference to one’s behaviour must be two separable steps.

If NARROW GUIDANCE is a condition on guidance, then the considerations appealed to by Glüer and Wikforss do seem to show that a truth norm cannot guide. It does not look plausible that we can respond to the no guidance argument by claiming that OUGHT can make a difference to what I believe after I’ve already formed a belief as to whether p. Steglich-Petersen accepts this (2013a, 282), and so do I.

WIDE GUIDANCE, on the other hand, does not require the guidance to involve two separable steps; it also allows for the guidance to be provided in just one step. This is because WIDE GUIDANCE allows that determining whether a norm’s antecedent condition is met can be itself part of the guidance that norm provides. It therefore allows that the guidance happens in one single step, rather than two separable ones. WIDE GUIDANCE does not, of course, rule out norms providing guidance in two separable steps. It just allows that they may also provide guidance in just one step.
The way in which Steglich-Petersen has outlined that OUGHT guides belief formation counts as guidance according to WIDE GUIDANCE for this very reason, i.e. because WIDE GUIDANCE allows that a norm can guide without that guidance being separable into two steps. If this is correct, then OUGHT can indeed provide genuine guidance by governing the way in which one forms a belief about whether its antecedent condition – i.e. whether p is true – is met. If one forms a belief as to whether p is true as part of an effort to determine whether p is true, then one counts as guided by OUGHT, and the fact that we can’t separate two steps doesn’t count against this fact.

If we are correct to accept WIDE GUIDANCE and reject NARROW GUIDANCE, then we can give answers to the two reasons which Glüer and Wikforss claimed made it strange to think of OUGHT as providing guidance.

Firstly, if WIDE GUIDANCE is correct, then the guidance provided by OUGHT doesn’t inevitably come too late. If I’m thinking about whether to believe that p, I don’t have to ‘wait’ until after I’ve formed a belief as to whether OUGHT’s antecedent condition before I can take guidance from the norm. OUGHT guides if I form a belief as to whether the norms’s antecedent condition is met as part of an effort to determine whether p is true. That guidance does not come after I’ve already answered the question I wanted guidance on – i.e. the question of whether to believe that p. Therefore, the guidance OUGHT provide does not come too late.

Secondly, if WIDE GUIDANCE correct, OUGHT does ‘make a difference’ to what I believe. It makes a difference because if I were to follow OUGHT I would end up with different beliefs than I would have if I were to follow some other norm, such as PLEASANT. Furthermore, preferring WIDE GUIDANCE to NARROW GUIDANCE means that it’s not the case that OUGHT has to make a difference to what I believe after I’ve formed a belief as to whether the norm’s antecedent condition is met. Rather it makes a difference to the way in which I form a belief as to whether the norm’s antecedent condition is met, i.e. to whether I do so as part of an effort to determine whether p is true (Steglich-Petersen 2013a, 283).

We should note here that being guided by OUGHT in the way Steglich-Petersen outlines doesn’t ensure that one conforms to OUGHT – i.e. that one has a true belief. But this doesn’t stand in the way of what Steglich-Petersen has outlined counting as guidance. This is because, as I outlined earlier, being guided by a norm is just for that norm to provide one with motivating reasons; being guided by the norm does not ensure
that one acts or believes in accordance with normative reasons.

In summary, the question of whether a truth norm can guide depends on what we take as necessary for a norm to provide guidance. If we assume narrow guidance, as Glüer and Wikforss do, then a truth norm can’t guide. If we reject narrow guidance and prefer wide guidance, then a truth norm can guide. And although Steglich-Petersen distinguishes these two conditions on guidance, he doesn’t give much in the way of reasons to reject narrow guidance and prefer wide guidance. He merely claims that Gluer and Wikforss’s assumption of narrow guidance is unmotivated (Steglich-Petersen 2013a, 282). Therefore, in order to bolster Steglich-Petersen’s response to Glüer and Wikforss, I’m now going to give some reasons to prefer wide guidance to narrow guidance.

3. In defence of Wide Guidance

I’m going to give two reasons to prefer wide guidance to narrow guidance. The first is negative, the second is positive. Firstly, I’ll respond to a counterargument from Glüer and Wikforss, and argue that they fail to give any reason to prefer narrow guidance to wide guidance. Secondly, I’ll argue that narrow guidance is too restrictive, because there are many instances of what intuitively looks like belief formation being guided by norms – norms other than a truth norm – which narrow guidance rules out. Therefore, we should prefer wide guidance.

3.1. Glüer and Wikforss’ counterargument: Wide Guidance is too wide

Glüer and Wikforss answer Steglich-Petersen’s response by arguing that the wide conception of guidance Steglich-Petersen has outlined is not a genuine sense of guidance, because it is too broad. This is because, they argue, wide guidance allows things that don’t look at all like guidance to count as guidance. They conclude that wide guidance should be rejected, and narrow guidance should be accepted. If this is correct, then the no guidance argument should still lead us to reject a truth norm.

The key cases which show wide guidance to be too broad, according to Glüer and Wikforss, are cases in which one’s following a norm causes one to be in accordance with that norm, but through some deviant causal chain. They illustrate this with the following example, in which someone’s acceptance of the alternative norm on belief pleasant makes a difference to their beliefs, but they don’t intuitively count as following
Imagine the following: a mad scientist has secretly (and wirelessly) hooked up S’s brain to a computer in such a way that if she accepts a norm making the correctness of believing p depend on property F, considers whether believing p would be F, and comes to the conclusion that it would not, her ability to believe p is (temporarily) blocked. In such a scenario, [PLEASANT], if accepted, influences S’s belief formation in the way amounting to [the wide conception of] guidance – but intuitively, S clearly does not follow [PLEASANT]. Even though accepting the truth norm instead of [PLEASANT] would result in a difference to S’s belief formation, the way this difference is generated is no longer recognizable as falling under any intuitive notion of rule-guidance. (Glüer and Wikforss 2015, 278)

In this example, S meets WIDE GUIDANCE, because their following of PLEASANT makes a difference to their believing. But S intuitively doesn’t look guided by PLEASANT. And on this basis Glüer and Wikforss claim that “[t]he proper conclusion … is not that our conception of guidance is too narrow but that Steglich-Petersen’s is too wide” (2015, 278).

Have Glüer and Wikforss given good reason to reject WIDE GUIDANCE and favour NARROW GUIDANCE, and thereby blocked Steglich-Petersen’s response? I don’t think they have. This is because, in summary, WIDE GUIDANCE is not intended to be an analysis of guidance, only a necessary condition on guidance. Therefore, the fact that meeting WIDE GUIDANCE is not sufficient for being guided, as in Glüer and Wikforss’s example, is irrelevant.

In discussing whether we should accept WIDE GUIDANCE or NARROW GUIDANCE, Steglich-Petersen is not providing an analysis of guidance, i.e. providing both necessary and sufficient conditions for being guided by a norm. Rather, WIDE GUIDANCE and NARROW GUIDANCE are just two candidates for necessary conditions on guidance. In particular, WIDE GUIDANCE and NARROW GUIDANCE are, as I outlined above, answers to a question concerning how many ‘steps’ being guided by a norm involves, i.e. whether it must involve forming a belief about whether the norm’s antecedent condition is met and the norm influencing one’s behaviour as two separable steps (as NARROW GUIDANCE claims) or whether these two elements of guidance can happen simultaneously (as WIDE GUIDANCE allows).

Neither WIDE GUIDANCE nor NARROW GUIDANCE is supposed to offer an analysis of guidance. This can be seen by their logical form, i.e. they are both of the form ‘A norm N guides only if p’. This can also be seen by the fact that both feature the notion
of ‘following’ a norm in the conditions themselves. Both of them are only going to be part of the full story about what it is to follow or to be guided by a norm (as opposed to merely being in accordance with a norm).

Therefore, since WIDE GUIDANCE is not intended to offer a sufficient condition on guidance, Glüer and Wikforss’s deviant causal chain case gives us no reason to reject it.

Moreover, it especially doesn’t give us reason to reject it in favour of accepting NARROW GUIDANCE. This is because NARROW GUIDANCE, if read as an analysis of guidance, is equally vulnerable to deviant causal chain cases. This is easiest to see if we recall the conditions’ phrasing:

WIDE GUIDANCE: A norm N of the form ‘Do X when in C’ can guide S’s behaviour with respect to X only if S following N can make a difference to S X-ing.

NARROW GUIDANCE: A norm N of the form ‘Do X when in C’ can guide S’s behaviour with respect to X only if S following N can make a difference to S X-ing after S has formed a belief as to whether C obtains.

In their example, Glüer and Wikforss’s mad scientist sets up S’s brain so that if S forms a belief that believing that p would not be pleasant, her ability to believe that p is temporarily blocked. Presumably, it’s possible both to imagine a case in which this blocking take place immediately, i.e. as soon as S forms a belief that believing that p would not be pleasant, and to imagine a case in which the blocking only takes place later, i.e. once S’s belief that believing that p would be pleasant is about to influence their behaviour. Since it’s a deviant causal chain, we can just move the causes to a later point in time at which they make equal trouble for NARROW GUIDANCE, if it’s (mistakenly) read as an analysis of guidance.

What this shows is that the kind of deviant causal chain case appealed to by Glüer and Wikforss is irrelevant to whether we should accept WIDE GUIDANCE or NARROW GUIDANCE. It would only be relevant if they were attempts at an analysis of guidance, whereas they are in fact candidates for necessary conditions on guidance. And even if they were attempts at analysis, deviant causal chain examples would make equal trouble for both. Deviant causal chain examples are, therefore, tangential to the question to which WIDE GUIDANCE and NARROW GUIDANCE are supposed to be answers, i.e. the question of whether guidance has to involve two separable steps. Glüer and Wikforss
have not said anything that should lead us to think we need to answer \textit{that} question by accepting NARROW GUIDANCE. Therefore, their counterargument to Steglich-Petersen’s response fails.

3.2. \textit{A reason against the two-step picture: it would rule out many other cases of belief formation being guided by norms}

Having dismissed Glüer and Wikforss’s objection to WIDE GUIDANCE, I’m now going to give a reason to reject NARROW GUIDANCE. We should reject NARROW GUIDANCE, I’ll argue, because it is far too restrictive. It is far too restrictive because it rules out many cases of what intuitively look like belief formation guided by norms, and would only allow beliefs to be guided by norms in the most exceptional circumstances. A conception of guidance that is this restrictive is, I suggest, implausible.

This can be seen if we focus on other norms on belief, ones much less controversial than a truth norm, such as norms enjoining us to avoid contradictory beliefs and to base our beliefs on evidence. The most common and everyday way in which such norms guide belief formation would not count as guidance according to NARROW GUIDANCE.

For example, consider the following norm:

\textsc{Avoid Contradiction}: If \(p\) is contradictory, one ought not to believe that \(p\).

If NARROW GUIDANCE is true, then in order for AVOID CONTRADICTION to guide, it would have to be able to make a difference to whether I believe that \(p\) after I’ve formed a belief as to whether \(p\) is contradictory.

Must being guided by AVOID CONTRADICTION in one’s belief formation or revision work like this? It doesn’t look like it has to.

To illustrate, imagine that I start off believing that God doesn’t exist. In attempting to follow AVOID CONTRADICTION, I think about whether the proposition that God doesn’t exist is contradictory. Say I conclude, perhaps on the basis of the ontological argument, that the proposition that God doesn’t exist is in fact contradictory. In concluding that it is contradictory, it seems plausible that, unless in exceptional cases, I will have already thereby revised my previous belief that God doesn’t exist.

This looks like a common and everyday way in which a norm like AVOID CONTRADICTION can guide belief revision. But if NARROW GUIDANCE is correct, this
does not count as guidance. It does not count as guidance, because it does not have the
two steps that NARROW GUIDANCE requires, i.e. AVOID CONTRADICTION has not made a
difference to what I believe after I’ve formed belief as to whether the proposition that
God exists is contradictory. Therefore, NARROW GUIDANCE is too restrictive, because it
rules out what looks like a common and everyday way in which a norm on belief guides.

The same seems to be the case with the following norm:

SUFFICIENT EVIDENCE: One ought to believe that p if and only if one has sufficient
evidence that p.

It is natural to think of this norm as, again, guiding belief formation and revision in a way
that does not involve the two steps that NARROW GUIDANCE requires. Say I start off not
believing that Brenda is thinking of leaving me. And say that, in attempting to follow
SUFFICIENT EVIDENCE, I think about whether I have sufficient evidence that Brenda is
thinking of leaving me. Thinking about her hesitance about moving in with me and the
fact that she keeps saying that she wants to go off travelling on her own, I conclude that
I do have sufficient evidence that Brenda is thinking of leaving me. Again, it seems
natural to think that in concluding that I have sufficient evidence, I will likely have
already thereby formed the belief that Brenda is thinking of leaving me.

Again this looks like a common and everyday way in which a norm like
SUFFICIENT EVIDENCE guides belief formation. And again NARROW GUIDANCE would
rule this out from counting as guidance. So NARROW GUIDANCE should be rejected as
too restrictive.

We should note here that in order to show that NARROW GUIDANCE is too
restrictive, we do not need to show that norms on belief never guide in the two step
fashion Glüer and Wikforss assume. In other words, we don’t need to show that is
impossible to conclude that p is contradictory or to conclude that one doesn’t have
sufficient evidence that p without thereby revising a prior belief that p. All we need to
show is that is possible for norms on belief to guide without two separable steps. That
suffices to show NARROW GUIDANCE is too restrictive.

In fact, I think it’s not only possible, but far more ubiquitous for norms on belief
to guide belief formation and revision without two separable steps. On the other hand,
cases in which there is some kind of gap or pause in between a) concluding that one has
sufficient evidence that p and b) forming a belief that p seem much more exceptional,
such as cases of temporary epistemic akrasia.

In this way, Glüer and Wikforss’s assumption of NARROW GUIDANCE is too restrictive. It doesn’t just rule out a truth norm from guiding belief formation; it also rules out many other instances of what look like belief formation being guided by norms. NARROW GUIDANCE, therefore, should be rejected, and the fact that a truth norm can’t guide according to NARROW GUIDANCE does not give us reason to reject a truth norm.

Given Glüer and Wikforss’s starting point, I don’t think we should be surprised that NARROW GUIDANCE turned out to be too restrictive. This is because the model of guidance they started with was not the way in which norms guide belief; their model was how norms guide action. This was evident in the fact that they drew their model of guidance from the way in which the practical norm ‘Buy low, sell high’ guides the actions of traders on the stock market.

The conclusion it seems right to draw here is that they have assumed that the kind of way norms guide action must hold for guidance across the board. But they haven’t provided any good reason to think that this must be the case. And it seems open to a defender of a truth norm to claim in response that belief and action are two very different things, and claim, as Engel does, that “[t]here is no reason to expect that epistemic norms regulate the formation and maintenance of our beliefs in the way practical norms do” (Engel 2013b, 53).

4. The non-centrality of guidance to the truth norm account of justification

In this chapter, I’ve defended Steglich-Petersen’s claim that if we adopt a wider, more inclusive conception of guidance than the one Glüer and Wikforss assume, then a truth norm like OUGHT can indeed guide belief formation. I’ve also argued that we should adopt this wider conception of guidance. Therefore, Glüer and Wikforss have not provided a reason to reject a truth norm.

I also want to end this chapter by noting the limited scope of this conclusion. This chapter only really concerns whether it’s possible to follow or be guided by a truth norm like OUGHT. Although I agree with Glüer and Wikforss that if a truth norm couldn’t guide we should reject the claim that belief is subject to a truth norm, my conclusion that a truth norm can guide has limited scope. This is because guidance is not central to the truth norm account of justification. It’s not central because the account should not involve the claim that being guided by OUGHT in one’s belief formation is
either necessary or sufficient for having a justified belief.

Firstly, being guided by OUGHT is not sufficient for having a justified belief because, as we saw earlier, following or being guided by a norm does not ensure believing in accordance with normative reasons. One may count as following OUGHT – if one forms a belief as to whether p as part of an effort of determining whether p is true – but still end up with a belief that’s unjustified – if one’s efforts are biased or sloppy – or a belief that’s merely false – if one’s efforts fail through no fault of one’s own. But such cases do not threaten the conclusion of this chapter. This is because this chapter only concerns guidance or motivation by norms. And in such cases one still counts as guided or motivated by OUGHT, even if one fails to conform to OUGHT or one fails to make a good attempt at conforming to OUGHT. In this way, being guided by OUGHT is not sufficient for justified belief.

Secondly, we also shouldn’t think that being guided by OUGHT is necessary for justified belief. This will take a bit longer to explain. It is because the truth norm account of justification is trying to explain why all justified beliefs are justified. And it’s implausible to think that for all justified beliefs, a subject has to follow or be guided by OUGHT with regards to that belief in order for it to be justified. Such an assumption need not and should not be part of the truth norm account of justification.

The explanation of justification in terms of a truth norm does appeal to the idea that one might follow one norm – e.g. a norm prescribing basing one’s beliefs on evidence – as a means to the end of following another norm – i.e. OUGHT. But this account should not claim that individual subjects actually do follow evidence norms as a means to the end of following OUGHT, in a way that reflected in their conscious or unconscious psychology. If such a claim were part of the truth norm account’s explanation of justification, that explanation would be highly implausible, because it’s not plausible that this means-end structure is reflected in the psychological of individual believers.

The claim that having justified beliefs is a good means to the end of conforming with OUGHT should instead be understood as an explanatory claim in normative epistemology, as a claim about how one norm, OUGHT, explains other norms, such as evidence norms and consistency norms. And it doesn’t seem that this explanation has to involve the means-end structure being reflected in a subject’s psychology.

This kind of explanation – i.e. one which a) claims that one norm is explained by
being a means to following another norm, but b) claims that this means-end relation does not need to figure in a subject’s psychology – is familiar from discussions of consequentialism in practical philosophy. It is often claimed that consequentialist norms should not be viewed as a ‘decision procedure’ but as a ‘criterion of rightness’. The classic statement of this view comes from Sidgwick:

[T]he doctrine that Universal Happiness is the ultimate standard must not be understood to imply that Universal Benevolence is the only right or always best motive of action. For, as we have before observed, it is not necessary that the end which gives the criterion of rightness should always be the end at which we consciously aim. (Sidgwick 1907, 413)

Sidgwick’s idea is that consequentialist norms don’t need to feature in an agent’s psychology for consequentialism to play a role in normative ethics of explaining what we ought to do. Viewed in this way, as a criterion of rightness, consequentialism can end up recommending seemingly un-consequentialist and easily-followable motives for action such as following rules of thumb, acting on the basis of virtuous dispositions, or acting from duty. We should, I suggest, view the truth norm’s explanation of subjective epistemic norms, like evidence norms and consistency norms, in comparable fashion. And in fact a number of defenders of truth norms have suggested we view a truth norm in this way (Engel 2013b, 42–43; Toppinen 2015, 399–400), as have some critics (Hattiangadi 2010, 427–28).

This means the truth norm account of justification does not have to rely on the (in my view dubious) assumption that a subject has to follow or be guided by a truth norm in order for their beliefs to be justified. Rather we should claim that it’s sufficient for a subject’s beliefs to be justified if they follow the subjective epistemic norms without an eye to a truth norm, so long as the normative reason why they should follow those subjective norms is because following them is a good means of conforming to a truth norm.\(^1\)

Does the fact that being guided by a truth norm is not necessary for being justified mean that the conclusion of this chapter is unnecessary? I don’t think so. If Glüer and Wikforss were right, and it were not possible for one’s belief formation to be guided by a truth norm, that alone, I believe, would be reason against the claim that

---

\(^1\) This material – and in particular the analogy with consequentialism – is developed from a paper written in collaboration with Christopher Cowie (Greenberg and Cowie 2016, sec. 7).
following subjective epistemic norms is a good means to the end of following a truth norm (how could norm A be a means to the end of following norm B if it were impossible to follow norm B?). Therefore, although following a truth norm directly plays a minimal role in the truth norm account of justification, answering Glüer and Wikforss is a necessary part of defending that explanation.

5. Conclusion

In this chapter, we’ve rejected the no guidance argument. It only would give us reason to reject a truth norm if we adopted an implausibly restrictive understanding of what it is for a norm to guide. However, we’ve also seen that a truth norm guiding is not centrally involved in the truth norm account of justification.

The real question of whether a truth norm can explain justification is whether the means one should take to conform to a truth norm actually would involve having beliefs that we pre-theoretically take to be justified. This is the question I’ll discuss for the remainder of this thesis.
SHOULD I BELIEVE ALL THE TRUTHS?

Should I believe something if and only if it’s true? Many philosophers have objected to this kind of truth norm, on the grounds that it’s not the case that one ought to believe all the truths. Philosophers who want to defend the claim that belief is subject to a truth norm often respond to this problem by reformulating a truth norm in ways that do not entail that one ought to believe all the truths. This chapter argues that that many of these attempts at reformulation have been missteps. A number of these different reformulations are incapable of carrying out a central role a truth norm is meant to play, that of explaining justification. The truth norm I’ll defend, however, avoids the problems associated with a prescription to believe all the truths, but doesn’t thereby fail to explain justification. This norm, introduced (but not defended) by Conor McHugh, states that if one has some doxastic attitude about p – i.e. if one believes, disbelieves, or suspends judgement about whether p – then one ought to believe that p if and only if p is true.

1. *The believing-all-the-truths problem*

Consider the following formulation of a truth norm:

**OUGHT**: One ought to believe that p if and only if p is true.

Because it is biconditional in form, **OUGHT** entails that one ought to believe all the truths. Many have claimed that this should make us reject **OUGHT**, and any truth norm that prescribes believing all the truths. We can outline three main reasons why an obligation to believe all the truths is problematic.

Firstly, an obligation to believe all the truths violates ‘**OUGHT**’ IMPLIES ‘**CAN**’:

‘**OUGHT**’ IMPLIES ‘**CAN**’: If one ought to φ, then it is possible for one to φ.

It’s not possible to believe all the truths. And, as Krister Bykvist and Anandi Hattiangadi have pointed out, there also seem to be truths that are too complex to grasp – such conjunction of all the truths – let alone to believe (2007, 279). And since ‘**OUGHT**’ IMPLIES ‘**CAN**’ is true, it cannot be the case that one ought to believe all the truths. And
so, it is concluded, OUGHT is clearly false (cf. Boghossian 2003, 37; Engel 2004, 82).

Secondly, there are some propositions that are logically impossible to truly believe, but which can be true. Bykvist and Hattiangadi, following Roy Sorensen (1988), call these propositions ‘blindspots’. They give the following example (Bykvist and Hattiangadi 2007, 281):

It is raining and nobody believes that it is raining.

Since this proposition can be true, when it is true, OUGHT implies that one ought to believe it. However, it is not truly believable. If I believe it, then it’s not true.

Bykvist and Hattiangadi argue that we should reject a norm like OUGHT which entails that one ought to believe true blindspot propositions. This is because, they claim, the obligation to believe true blindspot propositions is an obligation one cannot satisfy. As soon as one believes a blindspot, it’s then no longer the case that one ought to do so. Bykvist and Hattiangadi claim that we should reject a norm that entails these kinds of unsatisfiable obligations, because of the following principle:

‘OUGHT’ IMPLIES ‘CAN SATISFY’: If one ought to φ, then it is possible for one to φ while being obligated to φ.

Bykvist and Hattiangadi claim that this principle is just as plausible as ‘OUGHT’ IMPLIES ‘CAN’: “Just as one cannot be obligated to do the impossible, one cannot be obligated to satisfy requirements that are impossible to satisfy” (2007, 282). The possibility of true blindspots shows that OUGHT violates ‘OUGHT’ IMPLIES ‘CAN SATISFY’. And so Bykvist and Hattiangadi conclude that OUGHT is false.

Thirdly, that OUGHT prescribes believing all the truths seems problematic because some truths are trivial truths, such as, to use Daniel Whiting’s example, “all the truths about the length and colour of each hair on David Cameron’s left arm” (2012, 283). It doesn’t seem we ought to believe such trivial truths, and so, it is concluded again, OUGHT is clearly false.

And as well as being problematic in itself that OUGHT prescribes believing trivial truths, it also undermines a response one could give to the ‘OUGHT’ IMPLIES ‘CAN’ objection we sketched above. One might claim that although conforming to OUGHT is unachievable, OUGHT can still function as an epistemic ideal, i.e. a standard which we can at best approximate; but an ideal we can nevertheless appraise believers against (Engel
Trivial truths make this response look implausible, because those who fail to believe trivial truths don’t look like they are falling short even from the epistemic point of view. This is illustrated by the following example given by Ernest Sosa:

"Suppose you enter your dentist’s waiting room and find all the magazines missing. Deprived of your reading matter, you’re sure to doze off, but you need no sleep. Are you then rationally bound to reach for the telephone book in pursuit of truth? Were you not to do so, you would forfeit a chance to pluck some desired goods within easy reach. (E. Sosa 2001, 49)"

It seems absurd to suggest that in failing to memorize the phone book I am in any way falling short epistemically. So ought and its prescription that one believes all the truths, as well as not being achievable, doesn’t look like a good candidate for being an epistemic ideal (cf. Whiting 2012, 283–84).

The most common response to the believing-all-the-truths problem by defenders of a truth norm has been to reformulate the norm in ways that do not entail that one ought to believe all the truths. In fact, the vast majority of those who defend a truth norm defend formulations other than ought for this very reason (Boghossian 2003, 37; Engel 2004, 82; Whiting 2010, 216; McHugh 2012; 2014b; Wedgwood 2013b; Raleigh 2013).

Most of the reformulations of a truth norm so far given, I shall argue, have been missteps. In particular, many of these reformulated truth norms have implausible consequences that make them incapable of carrying out a central role a truth norm is meant to play, that of explaining justification.

Let’s recall how a truth norm like ought is supposed to explain justification. This is because one cannot follow a norm like ought directly, i.e. one can’t just tell whether one’s beliefs are true or compare what one believes with what is true. One is going to have to take some indirect means to try and ensure that one’s beliefs are in accordance with this norm. It seems that this would involve subsidiary norms, including, but not limited to, the following:

\begin{itemize}
  \item **Consistency**: One ought not to have inconsistent beliefs.
  \item **Evidence**: One ought to believe that p if and only if p is supported by the evidence.
\end{itemize}

Because one follows these subsidiary norms as a means to the end of following a truth
norm, that truth norm is claimed to be the *fundamental norm* on belief. This is the claim that belief is subject to evidence and consistency norms *because* belief is subject to a truth norm. It is then argued that beliefs that are in accord with these subsidiary evidence and consistency norms just are beliefs that we pre-theoretically take to be justified. As I outlined in Chapter 1, this is the way in which a truth norm is supposed to explain justification.

In §2 and §3, I’ll show that many truth norms that are reformulated so as to deal with the believing-all-the-truths problem have implausible consequences that make them incapable of being the fundamental norm on belief that explains justification. In §4, however, I’ll argue that there is one reformulated norm, originally put forward (but not defended) by Conor McHugh, which does not entail that one ought to believe all the truths, but doesn’t thereby become incapable of explaining justification.

2. Against permissive truth norms

A first way of reformulating a truth norm is by making it *permissive*, i.e. so that it doesn’t prescribe believing *all* the truths, it just prescribes believing *only* truths, or prohibits believing *falsehoods*. The following are examples of permissive truth norms:

**ONLY IF:** One ought to believe that p only if p is true. (Boghossian 2003, 37; Engel 2004, 82)

**MAY:** One may believe that p if and only if p is true. (Whiting 2010, 216)

**OUGHT NOT:** One ought not to believe that p if and only if p is false. (Raleigh 2013, 249)

These permissive truth norms don’t entail that one ought to believe all the truths, and they do not prescribe that one ought to believe the three problematic kinds of truths we discussed above: ungraspably complex truths, trivial truths, or true blindspots.

However, the problem with these permissive truth norms is that they don’t entail that one *ought* to believe *anything* (cf. Bykvist and Hattiangadi 2007, 280). Therefore, the best way of ensuring that one conforms to any of these norms is to suspend judgement about everything, and to adopt the pyrrhonian sceptical stance of *epoche*. This is because suspending judgement about whether p guarantees conformity with a permissive truth norm; believing that p, on the other hand, even if one has overwhelming evidence in
support of p, is needlessly risky from the point of view of conforming with a permissive truth norm. This means that a permissive truth norm cannot explain justification, because we wouldn’t pre-theoretically take someone who suspended judgement about everything to be justified in doing so (cf. McHugh 2012, 18).

It’s worth noting that this objection to permissive truth norms relies on the properties of being justified or unjustified being properties that can not only apply to beliefs, but also to failures to believe, i.e. the kind of failure to believe that is involved in suspending judgement. This is a sense of justification does seem to fall somewhere in between the standard distinction between having propositional justification (having justification to believe that p) and doxastic justification (being justified in believing that p). It’s like doxastic justification, because it seems to be an assessment of someone qua believer. But it is unlike doxastic justification, because being justified or unjustified in the doxastic sense is usually defined in terms of having a belief (see Turri 2010). However, I think it’s necessary in any case to allow that justification can apply to failures to believe; it’s necessary in order for our theory of epistemic justification to account for all kinds of epistemic ‘wrongdoing’, e.g. to account for the epistemic ‘wrong’ of excessive scepticism.

I’ll now discuss some possible responses to my rejection of a permissive truth norm. Firstly, one may respond that I have not shown that a permissive truth norm cannot explain justification; I’ve just shown that a permissive truth norm cannot explain justification by itself. It may explain justification together with other norms that do positively prescribe having some opinions in some cases.

I grant that I haven’t ruled out a permissive truth norm from playing some role in explaining justification together with other norms on belief. But that’s not what I was trying to do. Defenders of a permissive truth norm typically claim that it is the fundamental norm on belief (see e.g. Boghossian 2003, 38–39; 2005, 211–12; Engel 2004, 82; Whiting 2010, 220–22; Raleigh 2013, 256–57). And all I have been trying to show is that this claim is unsustainable, because the best means of conforming to a permissive truth norm is to suspend judgement about everything.

A second response one could give in defence of a permissive truth norm is given by Whiting. Whiting argues that a truth norm being permissive doesn’t stop it from explaining justification because, he claims, we should also think of subjective epistemic norms – i.e. the norms one follows as a means to the end of conforming with a truth norm – as being permissive (2013b, 129–31). In other words, we should claim that
following a permissive truth norm involves following a permissive evidential norm, such as the following:

**PERMISSIVE EVIDENCE**: One may believe that \( p \) if and only if one has sufficient evidence that \( p \).

This evidential norm is also permissive because it does *not* state that if one has sufficient evidence that \( p \), one ought to believe that \( p \). And Whiting claims we should think of evidential norms as permissive like this one – rather than requiring one to believe in any case in which one has sufficient evidence that \( p \) – because, he claims, there are infinitely many trivial propositions for which I have sufficient evidence. We can see this, Whiting claims, if we think about the logical consequences of propositions we do have sufficient evidence about. For example, if I have sufficient evidence that London is in England, then I also have sufficient evidence that London is in England or Tolstoy wrote *Great Expectations* (Whiting 2013b, 130). Being epistemically justified, Whiting claims, does not demand believing this second proposition, so we should think of evidential norms on belief as permissive (2013b, 130; cf. Nelson 2010; Littlejohn 2012, 46–47).

This response does not deal with the problem I’ve raised. Firstly, it doesn’t look like a permissive truth norm like MAY can in fact explain why belief is subject to even a permissive evidential norm like PERMISSIVE EVIDENCE. This is because, as I outlined earlier, a truth norm is supposed to explain why belief is subject to subsidiary norms, such as evidence and consistency norms, because conforming to those subsidiary norms, it is claimed, is a good means of ensuring conformity to a permissive truth norm. The problem for Whiting is that PERMISSIVE EVIDENCE is *not* a good means of ensuring conformity with MAY. Given that the best means of ensuring that one does not violate MAY is to suspend judgement about everything, being permitted to believe that \( p \) if and only if one has sufficient evidence that \( p \) (i.e. PERMISSIVE EVIDENCE) is *not* a good means of ensuring one doesn’t violate a permission to believe that \( p \) if and only if \( p \) is true (i.e. MAY). This is because suspending judgement about everything guarantees conformity with a norm like MAY which permits believing all and only the truths, whereas believing that \( p \) when one has sufficient evidence does *not* guarantee conformity with MAY (given the plausible fallibilist assumption that having sufficient evidence that \( p \) does not entail \( p \) is true). Therefore, from the point of view of conforming with MAY, PERMISSIVE EVIDENCE, is a needlessly risky permission. Therefore, PERMISSIVE EVIDENCE cannot be
explained by MAY.

Secondly, even if PERMISSIVE EVIDENCE could be explained by MAY, PERMISSIVE EVIDENCE still cannot explain justification, and for the very same reason that a permissive truth norm cannot explain justification. Because of its permissive nature – because it doesn’t entail that one ought to believe anything – the best way of following PERMISSIVE EVIDENCE is to suspend judgement about everything; that guarantees conformity with PERMISSIVE EVIDENCE. This means that a permissive evidential norm cannot explain justification, because, again, we wouldn’t take someone who suspended judgement about everything to be justified in doing so.

A third response one could give in defence of a permissive truth norm is given by Thomas Raleigh. Raleigh claims that someone who takes the pyrrhonian option of suspending judgement about p no matter how strongly the evidence supports p would in fact be a committing an error according to a permissive truth norm. This is because, he claims, suspending judgement about p no matter how strongly the evidence supports p is an attitude which “embodies a false belief about what the evidence supports” (Raleigh 2013, 263). Therefore, although the pyrrhonian does not have a false belief about p, he does have a false second order epistemological belief about the evidence in support of p. And in having that latter belief, he is violating the permissive truth norm.

I admit that someone who suspends judgement about p no matter how strongly the evidence supports p has a false belief about what the evidence supports. But it’s question begging to assume this in defence of a permissive truth norm. This is because if a permissive truth norm were the fundamental norm that explains justification, then the pyrrhonian’s belief about what the evidence supports would no longer be false.

To make this clear we need to be specific about which belief of the pyrrhonian’s is supposed to be false. It cannot be a belief that the evidence supports p to a certain degree – e.g. a belief that the evidence makes p 95% likely, for example – because the pyrrhonian can suspend judgement about p and retain that belief. It has to be the pyrrhonian’s belief that the evidence does not support p to a degree that’s sufficient to make one justified in believing that p. But if a permissive truth norm were the fundamental norm on belief, then that belief would not be false. If a permissive truth norm were the fundamental norm that explains justification, no amount of evidence in support of p could make one justified in believing that p, given that suspending judgement about p guarantees conformity with a permissive truth norm and believing p
does not. To assume that the pyrrhonian’s epistemological beliefs about epistemic support would be false, as Raleigh does, is question begging.

A fourth response a defender of a permissive truth norm could give is to claim that it’s impossible for one to suspend judgement about everything; one just has to have some beliefs in order to live one’s life (cf. Burnyeat 1980). A classic statement of this view comes from Hume:

The great subverter of Pyrrhonism or the excessive principles of scepticism is action, and employment, and the occupations of common life. (1748, 158–59)

Does this help the defender of the permissive truth norm? I don’t think it does.

This is because even if it’s true that one cannot suspend judgement about everything, that doesn’t deal with the fundamental problem a permissive norm has with explaining justification. This is because all that is strictly required for a permissive truth norm to fail to explain justification is for there to be some cases in which a) one would be justified in believing that p, and b) it’s possible for one to suspend judgement about p.

A permissive truth norm would fail to explain justification if there is any such case, because such a case would be one in which having a justified belief does not coincide with taking the best means to conform to a permissive truth norm. And it’s highly implausible to deny that there are any such cases, even if we grant the Humean point that it’s impossible to suspend judgement about everything. Therefore, the impossibility of suspending judgement about everything doesn’t touch my argument against permissive truth norms.

The lesson we should draw from this is that if we want to explain justification through norms on beliefs, that explanation is only going to work if not all norms on belief are permissive norms – i.e. norms that only prohibit having certain beliefs, but never positively prescribe believing anything. If there are only permissive norms on belief – whether they be truth norms, evidential norms, or something else – the best way of ensuring conformity with such norms will be take the pyrrhonian option of suspending judgement about everything (or about as much as one can); doing anything else will be needlessly risky.

3. Against Wedgwood’s conditional truth norm

A second way of reformulating a truth norm so that it doesn’t prescribe believing all the
truths is to make it conditional, i.e. a norm which only comes into force if certain other conditions are met (conditions other than truth, of course). The following norm, put forward by Ralph Wedgwood, is an example:

CONSIDERS: If one consciously considers p, then one ought to believe that p if and only if p is true. (Wedgwood 2013b, 127)¹

CONSIDERS, unlike OUGHT, does not entail that one ought to believe all the truths. This is because there are many true propositions which one does not consciously consider.

Does this mean that CONSIDERS does not prescribe believing the three problematic kinds of truth discussed above?

It certainly does mean that CONSIDERS doesn’t prescribe believing ungraspably complex truths. Presumably, if p is ungraspably complex, it’s also impossible for me to consciously consider p. And if that’s the case, then CONSIDERS will never prescribe that I believe an ungraspably complex truth. Therefore, CONSIDERS will not violate ‘OUGHT’ IMPLIES ‘CAN’ in the same way as OUGHT did.

Opting for CONSIDERS also seems to deal well with trivial truths. CONSIDERS means that I’ll never be obliged to believe any true propositions about the hairs on David Cameron’s arm or true propositions about the phone numbers of people I’ve never met, because I have never consciously considered those propositions and almost definitely never will.

However, CONSIDERS does not deal so well with blindspots. Let’s assume for the sake of argument that the proposition that it is raining and nobody believes that it is raining is true. According to CONSIDERS, if you consciously consider that proposition (as presumably you have done at least once while reading this chapter), then you ought to believe it (at the very least while you were considering it). That is an unsatisfiable

¹ Wedgwood doesn’t actually defend CONSIDERS. He defends the following truth norm, which doesn’t feature the concept ‘ought’, but instead features the concept ‘correct’:

CORRECT: A belief that p is correct if and only if p is true. (Wedgwood 2002, 272)

And although he doesn’t put it into the content of his truth norm, Wedgwood makes it clear that he regards CORRECT as only concerning propositions one consciously considers (2002, 273; 2008, 5, fn.8). For simplicity, I’m going to discuss CONSIDERS rather than CORRECT, because CONSIDERS is a more direct reformulation of the norm we started with, OUGHT.
obligation. CONSIDERS therefore violates ‘OUGHT’ IMPLIES ‘CAN SATISFY’, just as OUGHT does.

In order to get over this problem, Bykvist and Hattiangadi point out, we need to adopt the following modification of CONSIDERS:

TRULY BELIEVABLE: If one consciously considers p and p is truly believable, then one ought to believe that p if and only if p is true.

This norm does not entail that one ought to believe that it is raining and nobody believes that it is raining, even if it is true and one consciously considers it, because it’s not a proposition that’s truly believable.

However, Bykvist and Hattiangadi claim that this reformulation also fails to deal with blindspots. While TRULY BELIEVABLE admittedly does not entail that one ought to believe any true blindspots, Bykvist and Hattiangadi claim it is still unsatisfactory because it fails to entail that one ought not to believe true blindspots. And this makes TRULY BELIEVABLE remain problematic, they claim, because “intuitively, … precisely the right response to the proposition that it is raining and nobody believes that it is raining [is that] you should not believe it even if it is true” (Bykvist and Hattiangadi 2007, 282). Therefore, neither CONSIDERS nor TRULY BELIEVABLE, Bykvist and Hattiangadi conclude, can deal with blindspots.

We have seen that CONSIDERS, although it deals with ungraspably complex truths and trivial truths, still faces problems dealing with true blindspots. I also want to raise one more problem for CONSIDERS. CONSIDERS also fails to explain justification, much as the permissive truth norms I’ve already discussed failed to do so, though for rather different reasons.

The particular reason why CONSIDERS cannot explain justification is because it has implausible consequences in some situations in which one has overwhelming evidence that p but one fails to believe that p. Specifically, in such situations CONSIDERS has the result that if one does not consciously consider p, one cannot count as unjustified in failing to believe that p. This result, as I’ll now show, is implausible.

I’ll illustrate with an example. Say’s there’s a boss of a company, called Diane, who has overwhelming evidence that one of her employees has repeatedly abused his position. In particular, Diane carries out periodic reviews of the company’s accounts, and because of that has strong evidence that a particular employee has been secretly and
repeatedly spending company money himself. However, Diane avoids ever consciously considering whether her employee has abused his position for some pragmatic reasons: perhaps she avoids considering this because the employee’s a close friend and she can’t bear the thought that he’d steal from the company, or to avoid having to confront him about it, or to avoid being sued. And the fact that Diane, for these pragmatic reasons, avoids thinking about it means that she fails to believe what her evidence overwhelmingly indicates, i.e. that her employee has abused his position.

I take it that the intuitive response to such a case is that Diane is epistemically unjustified for failing to believe that her employee has abused his position. But if Diane is epistemically unjustified, then that creates a problem for CONSIDERS. In particular, it means that CONSIDERS cannot explain justification either.

Recall that truth norm like CONSIDERS is supposed to explain justification because taking the means that one should to conform to it with regards to p is supposed to result in one having a justified belief that p (or being justified in suspending judgement about whether p). Diane’s failure to believe that her employee has abused his position looks unjustified. But CONSIDERS cannot believe that her employee has abused his position, Diane has not taken any bad means of conforming to CONSIDERS. This is precisely because CONSIDERS does not obligle one to believe anything about p if one doesn’t consciously consider p; if one doesn’t consciously consider p, one is necessarily off the hook.

This case shows that CONSIDERS cannot explain justification because it’s a norm that makes justification too restrictive. If CONSIDERS were the fundamental truth norm that explains justification, then one could only be unjustified in relation to propositions one actually consciously considers. But the case of Diane shows this to be implausible, because it’s a case in which someone definitely looks unjustified in failing believe that p, but in relation to a proposition which that person does not consider. Therefore, CONSIDERS fails to explain justification as well.

A first objection a defender of CONSIDERS might make to this case would be to claim that Diane should have consciously considered whether her employee has abused her position, perhaps because of her fiduciary responsibility to the company’s shareholders. If this is the case, we can still claim that Diane has done something, but what she has done wrong is not explained by CONSIDERS, but by a distinct norm about what she ought to consciously consider.
I have two responses to this objection. Firstly, it seems to me that what Diane has done wrong is, at least in part, an epistemic failure. And if the only thing Diane has done wrong, according to the defender of CONSIDERS, is to violate a distinct norm about what she ought to consciously consider, then that doesn’t look like an epistemic failure; it looks more like a practical or moral failure. Claiming that Diane has only violated a norm about what she ought to consciously consider looks like an overly narrow picture of what kinds of failure occur in such a case.

Secondly, and more importantly, if the defence of CONSIDERS depends on appealing to distinct norms about what one ought to consciously consider, that means that CONSIDERS cannot be the fundamental norm on belief that explains justification, i.e. it cannot by itself explain why belief is subject to other epistemic norms. And just as was the case with the permissive truth norms, the key point I have intended to make is that CONSIDERS cannot be the fundamental norm that explains justification. Those who defend truth norms that only apply to propositions which one consciously considers, such as Wedgwood, do claim that such a truth norm can be the fundamental norm on belief that explains justification (Wedgwood 2002, 271–83). That claim is undermined if the defender of CONSIDERS has to appeal to other norms about what one ought to consciously consider in order to deal with cases like that of Diane.

A second way in which a defender of CONSIDERS might object to the Diane example involves being more specific about the tense of ‘considers’ in the norm. One may claim that the Diane example only is problematic for the following norm:

CONSIDERS-PRESENT: If one is currently consciously considering p, then one ought (at that time) to believe that p if and only if p is true.

But not for the following norm:

CONSIDERS-PAST: If one has consciously considered p at some time, then one ought (from then on) to believe that p if and only if p is true.

One may claim that, in the case I’ve described, it would be plausible that Diane has at some point consciously considered that whether her employee has abused his position if she is to count as having overwhelming evidence that he has done so. If that’s the case, then the case of Diane does not create problems for CONSIDERS-PAST. If Diane has consciously considered whether her employee has abused his position at some point,
then she cannot avoid CONSIDERS-PAST’s prescription to believe the truth about her employee by avoiding thinking about it, and the case does not create a problem for CONSIDERS-PAST’s explanation of justification.

I have two responses to this possible specification of CONSIDERS. Firstly, the claim that Diane must have at some point considered whether her employee had abused his position seems implausibly strong. One cannot support that claim by merely describing one version of my example in which Diane has at some point considered whether her employee had abused his position. Rather, to defend CONSIDERS-PAST one must show that it’s impossible for Diane to have never consciously thought about whether her employee has abused his position, but to also be unjustified in failing to believe that he has done so. This looks highly implausible. It seems at least possible for Diane, in the case I’ve described, to have never thought about whether her employee has abused his position, but nevertheless be unjustified in failing to believe that he has done so.

More generally it seems we’re familiar with cases like this, i.e. in which people are unjustified in failing to believe things that they have never consciously considered. Consider a case in which a member of a family has done something unspeakably horrible, which none of the other family members suspected. Also consider a case in which a wife decides to divorce her husband, who has never suspected that his wife was thinking of leaving him. People in cases like this sometimes do recognise, though usually only in retrospect, that there was in fact a good deal of evidence that supported the thing that they never suspected. They might say “the thought never crossed my mind that Henry could have done that [or: that Brenda was thinking of leaving me], but now I think about it, the writing was on the wall’. If this a true description of these cases, then they look like cases in which people are unjustified in failing to believe things that they had never consciously considered. And if this is the case, then CONSIDERS-PAST also fails to explain justification.

Secondly, even if it were true that Diane must have consciously considered whether her employee has abused his position at some point, opting for CONSIDERS-PAST over CONSIDERS-PRESENT in order to deal with cases like this is a bad move dialectically. In particular, the shift to CONSIDERS-PAST makes dealing with trivial truths more difficult. This is because CONSIDERS-PAST has the result that one ought to believe any trivial truth that one has consciously considered at some point.

To illustrate, take the proposition that there have been an even number of British
Prime Ministers. Having read the last sentence, presumably you have consciously considered the proposition that there has been an even number of British Prime Ministers (assuming that reading involves consciously considering). According to CONSIDERS-PAST, you now have the obligation to believe that there has been an even number of British Prime Ministers, since that’s true. And it seems that I get you to be obliged to believe any other trivial truth in a similar fashion. This makes CONSIDERS-PAST look like not much of an improvement over OUGHT with regards to trivial truths.

Moreover, what made CONSIDERS-PAST deal (comparatively) better with the case of Diane is precisely what makes it have this problem dealing with trivial truths. CONSIDERS-PRESENT deals with trivial truths in a (comparatively) better way. It seems more plausible that that one is obliged to believe a trivial proposition if and only if it’s true if that obligation only lasts as long as one actually is consciously considering that proposition. But this feature of CONSIDERS-PRESENT made it deal badly with cases like that of Diane, because it enables one to get off the hook if one avoids consciously considering the proposition in question.

In summary, reformulating a truth norm so that it only applies to propositions one consciously considers makes it incapable of carrying out the role of explaining justification. Such a reformulated truth norm has the result that one can only be unjustified in relation to propositions one actually consciously considers. But this, as the case of Diane shows, is implausible.

4. In defence of an alternative conditional truth norm

Although we’ve seen that CONSIDERS fails to explain justification, I want to suggest making a truth norm conditional is the right way to deal with the believing-all-the-truths problem. In particular, I’m going to argue that adopting the following conditional truth norm, discussed (but not defended) by Conor McHugh, is the right way to answer the believing-the-truths problem:

**DOXASTIC ATTITUDE**: If one has some doxastic attitude towards p, one ought to believe that p if and only if p is true. (McHugh 2012, 12)

---

2 Since the actual number is 76, if we start from Robert Walpole. There’s Theresa May, in addition to the 75 past Prime Ministers listed here: [https://www.gov.uk/government/history/past-prime-ministers](https://www.gov.uk/government/history/past-prime-ministers).
Having ‘some doxastic attitude towards p’ involves either believing that p, or disbelieving p (i.e. believing that not-p), or suspending judgement as to whether p. I’ll give more detail about the last of these three options shortly, and particular about why we should think of it as having a doxastic attitude.

How does opting for DOXASTIC ATTITUDE enable us to answer the believing-all-the-truths problem? Showing this involves two main steps.

Firstly, I’ll argue, in §4.1, that DOXASTIC ATTITUDE does not have the kind of problem CONSIDERS had with the case of Diane. Whereas it was implausible that one could be unjustified only in relation to propositions one actually consciously considers, it is plausible, I’ll suggest, that one can be unjustified only in relation to propositions one has some doxastic attitude towards. This means that DOXASTIC ATTITUDE doesn’t fail to explain justification in the way that CONSIDERS does.

Secondly, I’ll argue, in §4.2, that DOXASTIC ATTITUDE does not prescribe believing the three problematic kinds of truth discussed above: ungraspably complex truths, trivial truths, and true blindspots. I’ll admit that DOXASTIC ATTITUDE has some difficulty with dealing with trivial truths and true blindspots. This is because DOXASTIC ATTITUDE prescribes that one ought to believe some truths of these kinds because one can sometimes have some attitude towards them. But I’ll argue that this can be dealt with by appealing to the way in which a truth norm like DOXASTIC ATTITUDE explains justification. Firstly, I’ll argue that following DOXASTIC ATTITUDE will never involve believing blindspots. Secondly, although I’ll accept that following DOXASTIC ATTITUDE will involve believing some trivial truths, it does so in a way that’s unproblematic.

This second point is one we could also make, without changing too much, in defence of some other truth norms; in particular, one could make such a point in defence of CONSIDERS. Therefore, the key reason to adopt DOXASTIC ATTITUDE over the other truths norms I’ve discussed is that it provides an answer believing-all-the-truths problem that does not thereby make it incapable of explaining justification.

4.1. Why DOXASTIC ATTITUDE does not make justification overly restrictive

I’ll first explain why DOXASTIC ATTITUDE does not have the kind of problem CONSIDERS had with the case of Diane. The problem with CONSIDERS was that if it were the fundamental norm explaining justification, then justification would be overly restrictive; in particular, it would have the result that one could only be unjustified in relation to
propositions one consciously considers. But we saw this to be implausible. One can, as the case of Diane showed, be unjustified in failing to believe that p even if one does not consciously consider p.

DOXASTIC ATTITUDE does not, I suggest, have an analogous problem. It’s much more plausible, I’ll argue, that one in fact can only count as being unjustified in failing to believe that p if one counts as having some doxastic attitude about whether p, i.e. if one counts as either believing, disbelieving, or suspending judgement about whether p (where suspending judgement counts as having a distinct doxastic attitude towards p which is distinct from merely neither believing nor disbelieving p).

Why should we think that one can only have an unjustified belief that p if one has some attitude about whether p? The key reason, in summary, is that in cases in which someone intuitively looks unjustified for failing to believe that p, it’s only plausible that that person is unjustified if they also have some other doxastic attitude towards p.

We can see this, I will argue, if we think about what would be involved someone lacking any doxastic attitude about p, i.e. to not only neither believe nor disbelieve p, but to not even count as suspending judgement about whether p. Once we take into account what would be really be involved in someone not having any doxastic attitude about whether p, it is no longer plausible that in such a case that person could count as unjustified for failing to believe that p. If this is correct, then it is plausible that one can only count as unjustified in relation to propositions one has some doxastic attitude towards. And if that is the case, then DOXASTIC ATTITUDE does not result in an overly restrictive account of justification in the way that CONSIDERS did.

In order to show this, I’ll need to briefly discuss the reasons for thinking that suspension of judgement is a distinctive doxastic attitude.

It’s plausible that to suspend judgement about whether p is not just to neither believe nor disbelieve p. This can be illustrated by the fact that, to use Wedgwood’s example, even rocks and numbers count as neither believing nor disbelieving propositions, but they do not count as suspending judgement about them (2002, 272). Rather, suspending judgement about whether p seems to involve having some attitude about whether p, but just an attitude of neutrality rather than belief (though it needn’t involve being exactly neutral as to whether p, i.e. thinking that p is just as likely as not-p). In this vein, Scott Sturgeon describes suspension of judgement as committed neutrality (2010, 136).
There is a further question, which I won’t go into, about what constitutes this attitude of committed neutrality about whether p. Some claim that it is reducible to some other doxastic attitude. This could be a middling credence/confidence as to whether p (Christensen 2009, 757), or a middling confidence spread (otherwise known as a middling ‘mushy credence’) (Sturgeon 2010). It could also be a higher order belief, such as the belief that the evidence doesn’t support having a belief about whether p (Raleigh 2013, 263). Others claim that suspending judgement about whether p is a distinct and sui generis doxastic attitude towards p (J. Friedman 2013).

I’m going to set this question aside, because I don’t want to discuss which particular doxastic attitude suspending judgement should be analysed as (if any). Rather, I want to discuss what differentiates cases in which I suspend judgement about p – i.e. cases in which I have the distinct attitude of neutrality about whether p (however that is to be analysed) – from cases in which I merely lack belief about whether p. This is because I want to claim that all cases in which someone counts as unjustified in failing to believe that p (in which they don’t disbelieve p) are cases in which one has suspended judgement about whether p.

So what differentiates someone who suspends judgement about whether p from someone who merely lacks belief about whether p? A good guide to when someone counts as suspending judgement, I suggest, is if we need to appeal in any way to an attitude of neutrality about whether p as part of their cognitive economy, e.g. to explain why they acted in a certain way or had other attitudes. For example, say a holidaymaker was planning on going on holiday after the Brexit referendum, and changed half of their currency before the referendum and waited to change the rest after the referendum. In order to explain this holidaymaker’s behaviour, we need to appeal to attitude of neutrality about who would win the referendum.

On the other hand, there are other cases in which people neither believe nor disbelieve p where we don’t need to appeal to an attitude of neutrality about whether p. Cases in which one lacks the requisite concepts involved in p are clear examples of this. For example, most young children presumably neither believe nor disbelieve that Britain will enter a recession after Brexit. But since they lack the requisite economic concepts, we don’t need to appeal to an attitude of neutrality about whether there’ll be a recession as part of their cognitive economy, and so they don’t count as suspending judgement about this.
Other examples include cases in which one can’t fully grasp p because one’s not sufficiently acquainted with the content of p. For example, presumably I count as neither believing nor disbelieving that a particular hair on David Cameron’s left arm is exactly 1cm long, because I’m not sufficiently acquainted with the particular hairs on David Cameron’s arm. But because I can’t really grasp the content of the proposition in question, we don’t need to appeal to an attitude of neutrality as part of my cognitive economy. Therefore, although I neither believe nor disbelieve this proposition, I don’t count as suspending judgement about it either.

This is a rough sketch of the kind of factors which differentiate suspension of judgment from mere non-belief (though for more detail see J. Friedman 2013). But this rough sketch will be sufficient to defend the claim that one can that only be unjustified in failing to believe that p if one has some attitude towards p, i.e. either disbelief or suspension of judgement.

This claim is plausible because in cases in which someone intuitively looks unjustified for failing to believe that p, I think it’s only plausible that that person is unjustified if they also have some other doxastic attitude towards p. This can be illustrated with the case of Diane, but, as we’ll see, the point generalises.

In the case of Diane, at least as I described it above, I think it’s plausible that Diane counts some doxastic attitude about whether her employee has abused his position, whether that be a negative attitude of disbelief or a neutral attitude of suspension of judgement. It seems, as least in the case as I described it above, we would need to appeal to some doxastic attitude about whether her employee has abused his position – whether it be disbelief or an attitude of neutrality – as part of Diane’s cognitive economy in order to explain her actions and attitudes, e.g. to explain why she hasn’t sacked her employee, why she’s not worried about getting sued, etc.

And, more importantly, I think it’s plausible that Diane counts as unjustified only given the assumption that she has some doxastic attitude about whether her employee has abused his position. This can be shown by thinking about what would have to be different in order for it to be plausible that Diane counts as not having any doxastic attitude about whether her employee has abused his position, i.e. to not even count as suspending judgement about it. If we think about what would have to be different for Diane to count as not having any doxastic attitude about it, then it’s no longer plausible that in such a case Diane would count as unjustified.
What would have to be the case for Diane to lack any doxastic attitude about whether her employee abused his position? Recall the kinds of case I discussed earlier in which people neither believe nor disbelieve p, but don’t count as suspending judgement about whether p. The examples of cases of this kind I gave included cases in which people lack the requisite concepts involved in p and cases in which people aren’t sufficiently acquainted with p’s content. Because we don’t need to appeal to an attitude of neutrality as to whether p as part of their cognitive economy, such people plausibly count as not having any attitude towards p at all.

We could modify the case of Diane to be more like these cases. For example, let’s suppose that Diane doesn’t fully understand what it is for someone to abuse their position; perhaps she doesn’t have the concept of legal or moral obligation, so she doesn’t understand that her employees have certain duties in virtue of having the roles that they do.

I admit that it would be plausible that in such a case Diane would count as not having any attitude about whether her employee has abused her position, not even an attitude of neutrality. But I don’t think it’s plausible that in such circumstance Diane would count as unjustified in failing to believe that her employee has abused his position, even if she still has overwhelming evidence that he has. It doesn’t seem plausible that someone can count as unjustified for failing to believe that p if they lack the requisite concepts to understand p.

We can illustrate with another example. Consider a child whose mother is having an affair, but since the child doesn’t fully understand how adult relationships work, they don’t really understand what it is for their mother to have an affair. While I think it’s plausible that this child could count as having lots of evidence that their mother is having an affair (though they may only recognise in retrospect that it was evidence), I don’t think we would think such a child as being unjustified for failing to believe this. And likewise if Diane genuinely lacked the concepts involved in understanding what it is for employee to abuse his position, then I don’t think it’s plausible that she is unjustified for failing to believe that he has done so.

Both of these cases don’t plausibly look like cases of people being unjustified; they look more like cases of non-culpable ignorance. Therefore, changing the case in this way so that Diane counts as not having any doxastic attitude about whether her employee has abused his position makes it into a case in which it’s no longer plausible
that she’s unjustified.

This is, of course, just one way in which the case of Diane could be changed so that she does not count as having any attitude towards whether her employee has abused his position. And more generally the case of Diane, as I originally described it, is just one example of someone who looks unjustified in failing to believe that p. But I think there is a principled reason for thinking that we can extend the conclusions we’ve made about this case more generally, and conclude that one can only be unjustified in failing to believing that p if one has some doxastic attitude towards p.

This is because in order to make it the case that a subject doesn’t just suspend judgement about whether p, but doesn’t have any attitude about it at all, you need to invoke some sort of gap or distance between the subject and the proposition in question. That is required to support the claim that the subject doesn’t have any doxastic attitude about whether p, i.e. the claim that proposition in question plays no role in the subject’s (doxastic) mental life.

This was evident in my modified Diane example, in which I needed to invoke the possibility of Diane lacking the requisite concepts in order to create the right kind of gap in order for Diane to count as not having any attitude about whether her employee abused his position. But the existence of such a gap or distance between the subject and the proposition is also likely make it implausible that the subject could count as unjustified for failing to believe that p, i.e. it’s likely to make it into a case of non-culpable ignorance.

It doesn’t, therefore, look plausible that there will be any cases in which someone both plausibly looks like they don’t have any doxastic attitude about whether p and plausibly looks unjustified in failing to believe that p.

This should lead us to conclude, I suggest, that one can only be unjustified in failing to believe that p if one has some doxastic attitude about whether p. And if this is case, DOXASTIC ATTITUDE does not have an analogous problem to CONSIDERS. CONSIDERS made justification overly restrictive. If CONSIDERS were the fundamental norm that explains justification, that would mean that one could only be unjustified in failing to believe that p if one consciously considered p. The case of Diane, as I originally described it, showed this to be false. On the other hand, if DOXASTIC ATTITUDE is the fundamental norm that explains justification, that would mean that one could only be unjustified in failing to believe that p if one has some doxastic attitude about whether p.
But this, as I’ve just shown, is plausibly true. Therefore, DOXASTIC ATTITUDE does not result in an overly restrictive account of justification.

4.2. How DOXASTIC ATTITUDE answers the believing-all-the-truths problem

We can now move on to saying how DOXASTIC ATTITUDE deals with the believing-all-the-truths-problem. DOXASTIC ATTITUDE clearly doesn’t entail that one ought to believe all the truths because there are many truths which one doesn’t have any doxastic attitude about. But does it entail that one ought to believe the three problematic truths that we covered earlier: ungraspably complex truths, trivial truths and true blindspots?

On the one hand, it seems to deal easily with ungraspably complex truths. Presumably if p is ungraspably complex, it’s impossible for me to have any doxastic attitude about whether p. So DOXASTIC ATTITUDE will never prescribe believing ungraspably complex truths. Therefore, DOXASTIC ATTITUDE does not violate ‘OUGHT’ IMPLIES ‘CAN’ in the way that OUGHT did.

On the other hand, it might seem at first glance to have some trouble with dealing with trivial truths and true blindspots.

Let’s start with trivial truths. There are many examples of trivial truths which DOXASTIC ATTITUDE does not and will never prescribe believing, such as a truth about the length of a hair on David Cameron’s arm, because I don’t have any doxastic attitude towards this truth, and almost certainly never will.

However, it looks like DOXASTIC ATTITUDE might prescribe believing some trivial truths. This is because I might count as suspending judgment about some trivial truths. For example, say I’m home for Christmas, and my brothers insist on playing a board game, but the only game we have is an out of date edition of Trivial Pursuit. During this game, my brother asks me what the nationality was of the bronze medal winner in men’s 100m breaststroke at the Seoul Olympics. I tell him I don’t know the answer, and we proceed with the game.

The true proposition about the nationality of the bronze medal winner of the men’s 100m breaststroke shooting at the Seoul Olympics looks just as trivial as true proposition about a hair on David Cameron’s arm, and I look just as justified in having no belief about it. But I do seem to have a doxastic attitude about the latter proposition, namely suspension of judgement, because we would need to appeal to an attitude of neutrality about whether this proposition is true in order to explain by behaviour, i.e. my
telling my brother I didn’t know the answer. So according to DOXASTIC ATTITUDE, I ought to believe it. But an obligation to have a true belief about the nationality of the bronze medal winner at the men’s 100m breaststroke at the Seoul Olympics looks to be the kind of implausible obligation we were trying to avoid.

Secondly, DOXASTIC ATTITUDE seems to have some trouble dealing with true blindspots. Presumably I count as either disbelieving or suspending judgement about whether it’s raining but nobody believes it’s raining. But it also seems possible that it’s true that it’s raining but nobody believes that it’s raining, given that I’m not infallible. And if it’s true, then, according to DOXASTIC ATTITUDE, I ought to believe it. So it seems that DOXASTIC ATTITUDE prescribes believing some true blindspots.

The way to deal with these problems is not, I suggest, to further reformulate DOXASTIC ATTITUDE so that it does not prescribe believing any trivial truths or true blindspots. The better way to deal with these problems, I will argue, involves appealing to the way in which a truth norm like DOXASTIC ATTITUDE is supposed to explain justification. Recall that a truth norm like DOXASTIC ATTITUDE is supposed to explain justification because one should follow subsidiary norms as a means to the end of conforming to it. This enables us to show that the fact DOXASTIC ATTITUDE prescribes believing some true blindspots is unproblematic.

This is because although DOXASTIC ATTITUDE does prescribe believing some of these problematic truths, the means one should take to the end of conforming with DOXASTIC ATTITUDE will, I’ll argue, never involve believing true blindspots, and won’t prescribe believing any problematic trivial truths. The means may involve, for some people in some cases, believing trivial truths. But these cases, I’ll argue, look unproblematic. Therefore, the fact that DOXASTIC ATTITUDE prescribes believing such truths is, in the end, unproblematic. It is much easier and clearer to make this point with regards to true blindspots, so I’ll deal with them first in §4.2.1. I’ll then deal with trivial truths in §4.2.2.

Before that, however, I just want to reiterate that this way of dealing with the believing-all-the-truths problem is not exclusive to DOXASTIC ATTITUDE. One could equally appeal to way OUGHT, CONSIDERS, or MAY is supposed to explain justification, i.e. one could argue that the means one should take to follow those norms won’t involve believing the problematic kinds of truths we’ve been discussing (though this would be trivially the case with regards to permissive truth norms, because the means one should take to follow a permissive truth norm will never involve believing anything). The key
reason to favour DOXASTIC ATTITUDE over these other norms is that it can provide this kind of answer to the believing all the truths problem, but without suffering the problems the other norms had, i.e. without, as OUGHT did, violating ‘OUGHT IMPLIES ‘CAN’; without, as permissive truth norms did, entailing that pyrrhonian scepticism is true; without, as CONSIDERS did, making justification overly restrictive.

4.2.1. Why following DOXASTIC ATTITUDE will involve never believing blindspots

Let’s assume that I have some attitude about the proposition that it’s raining and nobody believes that it is raining, i.e. that I count as either disbelieving or suspending judgement about it. If this is the case, and if the proposition is true, then DOXASTIC ATTITUDE does prescribe that I ought to believe it.

This is unproblematic, I suggest, because the means one should take in order to conform to DOXASTIC ATTITUDE will never involve believing the proposition that it is raining and nobody believes that it is raining. And this is the case whether or not the proposition is true.

This is because if one does believe that it’s raining and nobody believes that it’s raining, one is guaranteed to fail in conforming with DOXASTIC ATTITUDE. This is because as soon as one believes it, it becomes false, and it thereby becomes the case that one has failed to conform to DOXASTIC ATTITUDE. And, of course, this is the case whether or not this proposition was true before one believed it.

On the other hand, if one does not believe that it is raining and nobody believes that it’s raining – by either disbelieving or suspending judgement about it – then one is not guaranteed to fail in conforming with DOXASTIC ATTITUDE. If it does turn out to be true that it is raining and nobody believes that it is, then in failing to believe this proposition one will fail to conform to DOXASTIC ATTITUDE. But it could be false that it’s raining and nobody believes that it is. And in that case one has succeeded in conforming to DOXASTIC ATTITUDE, because one has avoided having a false belief about whether it’s raining and nobody believes it is.

Therefore, clearly the best strategy from the point of view of conforming with DOXASTIC ATTITUDE is to not believe a blindspot proposition, as can be seen in the table below:
As the table shows, the option of believing a blindspot is guaranteed to result in failing to conform with DOXASTIC ATTITUDE, whereas the option of not believing a blindspot could possibly succeed. Therefore, a believer trying to conform with DOXASTIC ATTITUDE clearly should not believe any blindspots.

Furthermore, we can be sure of this in advance, because it’s written into the content of a blindspot. I don’t need to try to conform to DOXASTIC ATTITUDE by believing a blindspot, and then realise, much to my chagrin, that I’ve failed only after I’ve formed a belief.

Therefore, although DOXASTIC ATTITUDE prescribes believing some true blindspots, the means one should take to conform to DOXASTIC ATTITUDE will never involve believing blindspots, even if they are true.

Therefore, this is a way in which we can do justice to Bykvist and Hattiangadi’s intuition that the right response to propositions of this kind is that one shouldn’t believe them even if they are true. That’s not, admittedly, what DOXASTIC ATTITUDE prescribes directly, but it is prescribed by the subsidiary norms one should follow as a means to the end of DOXASTIC ATTITUDE.

There is still the fact that DOXASTIC ATTITUDE violates Bykvist and Hattiangadi’s principle ‘OUGHT’ IMPLIES ‘CAN SATISFY’. However, now we’ve shown that the means one should take to conform to DOXASTIC ATTITUDE will never involve believing blindspots, it’s unclear that there remains any theoretical support for ruling out

<table>
<thead>
<tr>
<th>Blindspot Options</th>
<th>Believing that it’s raining and nobody believes that it’s raining</th>
<th>Not believing that it’s raining and nobody believes that it’s raining</th>
</tr>
</thead>
<tbody>
<tr>
<td>It’s true that it’s raining and nobody believes it’s raining</td>
<td>Impossible</td>
<td>Failure to conform to DOXASTIC ATTITUDE (failure to have a true belief)</td>
</tr>
<tr>
<td>It’s false that it’s raining and nobody believes it’s raining</td>
<td>Failure to conform to DOXASTIC ATTITUDE (false belief)</td>
<td>Success in conforming to DOXASTIC ATTITUDE (avoidance of false belief)</td>
</tr>
</tbody>
</table>
DOXASTIC ATTITUDE on the grounds that it violates this principle.

The motivations for an obligation-limiting principle like ‘OUGHT IMPLIES ‘CAN SATISFY’ – assuming that it has the same kind of motivations at the more familiar obligation-limiting principle ‘OUGHT IMPLIES ‘CAN’ – typically are that we need such a principle in order rule out certain norms, e.g. to rule out norms that are over-demanding, norms that are unfair, norms that cannot provide guidance, or norms that lead one on pointless errands (see Andrić 2015 for a summary). It’s hard to see how any of these rationales can provide theoretical support for appealing to ‘OUGHT IMPLIES ‘CAN SATISFY’ in order to rule out DOXASTIC ATTITUDE, given that one can know in advance that the means one should take to follow DOXASTIC ATTITUDE will never involve believing a blindspot.

Therefore, the fact that it violates Bykvist and Hattiangadi’s principle ‘OUGHT IMPLIES ‘CAN SATISFY’ doesn’t provide us with reason to reject DOXASTIC ATTITUDE, because it’s not clear that the principle itself has sufficient theoretical support.

4.2.2. Why following DOXASTIC ATTITUDE will not involve believing (any problematic) trivial truths

The way DOXASTIC ATTITUDE allows us to deal with trivial truths is not so simple. Following DOXASTIC ATTITUDE will not prescribe believing the vast majority of trivial truths, but it will, I think, prescribe believing some of them. However, I’ll argue that this is unproblematic. The cases in which following DOXASTIC ATTITUDE does involve believing a trivial truth are cases in which one would look unjustified if one didn’t believe the trivial truth in question.

Explaining this requires us to be more specific about the particular subsidiary norms one should follow as a means to the end of conforming to DOXASTIC ATTITUDE. In particular, we need to claim one of the subsidiary norms is the following norm: SUFFICIENT EVIDENCE-DA: If one has some doxastic attitude about whether p, one ought to believe that p if and only if one has sufficient evidence that p.

---

3 The ‘DA’ indicates ‘doxastic attitude’, and serves to distinguish SUFFICIENT EVIDENCE-DA from SUFFICIENT EVIDENCE, which I introduced in Chapter 3. The only difference is that SUFFICIENT EVIDENCE-DA is conditional on having some doxastic attitude whereas SUFFICIENT EVIDENCE is not. I’ll
What does it mean to have ‘sufficient evidence’? It is supposed to express an intuitive requirement on justification. In order to have a justified belief that p, it is not enough for one to have *some* evidence that supports p. For example, say it’s midsummer, and I wake up and haven’t drawn the curtains yet. In that case, it being midsummer is some evidence that it’s not raining outside. But that’s not enough for me to have a justified belief that it’s not raining outside. That requires me to have evidence that supports the fact that it’s raining above a certain *threshold* that counts as *sufficient*, which in this case may just involve opening the curtains. This is the requirement that is encoded in **SUFFICIENT EVIDENCE-DA**.

Why should one follow **SUFFICIENT EVIDENCE-DA** as means to the end of conforming with **DOXASTIC ATTITUDE**? In brief, this is because **DOXASTIC ATTITUDE** prescribes *both* believing the truth *and* avoiding false beliefs. In order to strike a balance between these achieving these dual competing aims, one would need to follow a subsidiary norm that did two things. Firstly, it would need to prescribe believing that p *if* one has evidence that counts in favour of p’s truth, in order to give oneself a good chance of believing some truths. Secondly, it would need to prescribe believing that p *only if* one’s evidence supports p above a certain threshold, in order to give oneself a good chance of avoiding false beliefs. And **SUFFICIENT EVIDENCE-DA** is a norm that does both of these things. So one should follow **SUFFICIENT EVIDENCE-DA** as a means to the end of conforming to **DOXASTIC ATTITUDE**.

A full defence of a truth norm account of justification will include a more detailed explanation of when one’s evidence is sufficient, and a fuller defence of the claim that one should follow **SUFFICIENT EVIDENCE-DA** in order to conform to **DOXASTIC ATTITUDE**. I’ll give much more detail on these fronts in Chapter 6. But I’ve given as much detail as is needed in order to answer the problem of this chapter, the believing-all-the-truths problem, and in particular to deal with trivial truths.

This is because, first, the assumption that one should follow a norm that prescribes believing on the basis of sufficient evidence as a means to the end of following a truth norm is *dialectically* appropriate from the point of view of answering the believing-all-the-truths-problem. For if this assumption is *false*, then a truth norm won’t only discuss **SUFFICIENT EVIDENCE-DA** in this chapter, because this is the only chapter where it is relevant to have an evidence norm that is restricted to propositions one has some doxastic attitude towards.
be able to explain justification, but for reasons that are independent of the believing-all-the-truths problem.

And second, we don’t need to go into any detail about how much one’s evidence needs to support \( p \) in order to be above the threshold for sufficient evidence, and this because of the amount of evidence we typically have about trivial truths. Trivial truths typically are truths which we have an extremely small amount of relevant evidence about, if we have any at all. Therefore it’s clear that we’ll be below the threshold of sufficient evidence with regards to trivial truths, and we can know this without going into detail about how that threshold is set.

We can illustrate this with the example of the true proposition about the nationality of the bronze medal winner in the men’s 100m breaststroke at the Seoul Olympics. As I granted above, because I have an attitude about this proposition (suspension of judgment), DOXASTIC ATTITUDE entails I ought to believe the truth about this proposition. But following SUFFICIENT EVIDENCE-DA does not involve believing the truth about this proposition. This is because while I probably count as having some relevant evidence about this proposition (e.g. I know he’s more likely to be from a historically successful Olympic nation like the USA or the Soviet Union than, say, Mongolia or Luxembourg), I clearly don’t have anywhere near sufficient evidence about it. And this is the case regardless of how we determine the threshold for sufficient evidence.

Therefore, although DOXASTIC ATTITUDE does entail that I ought to believe the truth about the nationality of this medal winner, the means I should take to conform to DOXASTIC ATTITUDE – i.e. following SUFFICIENT EVIDENCE-DA – will not lead to believing this truth.

At this point, one may object that there is an easy route to getting sufficient evidence about the nationality of the bronze medal winner at the Seoul Olympics – e.g. by looking on Wikipedia. This may lead one to question whether I have in fact shown that following DOXASTIC ATTITUDE will not involve believing this truth. If I’m trying to ensure conformity with DOXASTIC ATTITUDE with regards to the truth about the nationality of the bronze medal winner, wouldn’t the best way to do that be to go on Wikipedia, rather than to suspend judgement? Doesn’t DOXASTIC ATTITUDE therefore mean that I ought to waste my time looking this up on Wikipedia?

I don’t think there is a problem here. The subsidiary norms one must follow as a means of conforming to DOXASTIC ATTITUDE I’ve been thinking about are all norms that
prescribe what one ought to believe. For this objection to succeed, we have to grant that there are also some subsidiary norms that prescribe one to perform certain actions. But it seems we can restrict the norms that a truth norm is supposed to explain solely to norms on belief.

Some have claimed that a truth norm should also prescribe some actions, e.g. that one ought to gather evidence (Horwich 2013, 29–31). But it seems acceptable to me if a truth norm only explains what one ought to believe in order to have justified beliefs, i.e. that a truth norm only explains when someone’s attitude towards p counts as justified or unjustified given the evidence concerning p they currently have. And for DOXASTIC ATTITUDE to explain that, it seems sufficient if the subsidiary norms one must follow as a means of conforming it are restricted to norms on belief. If there are any epistemic norms on action, they can be explained by other norms (norms on inquiry, perhaps).

Therefore, once we recognise that DOXASTIC ATTITUDE is only supposed to explain what one ought to believe in order to be justified, it then becomes clear that following DOXASTIC ATTITUDE will not involve looking up who was the bronze medal winner at the Seoul Olympics on Wikipedia. Rather, it will involve suspending judgement about it, because one does not have sufficient evidence.

This is just one example, but it seems we’ll be able to say the same about the vast majority of trivial truths. At least with all the canonical examples – e.g. truths about phone numbers, truths about grains of sand, etc. – it seems that I’m bound to not have sufficient evidence about such truths. Therefore, if for some reason I happen to have some doxastic attitude about such a truth, although DOXASTIC ATTITUDE will then entail that I ought to believe it, we can deal with it in the same way as the example above.

Therefore, although DOXASTIC ATTITUDE may prescribe believing some trivial truths, the means one should take to conform to DOXASTIC ATTITUDE will not involve believing the typical examples of trivial truths. This makes the fact that DOXASTIC ATTITUDE prescribes believing such truths unproblematic.

What about trivial truths which one both has an attitude about, and has sufficient evidence for? Doesn’t conforming to DOXASTIC ATTITUDE involve believing those truths?

I grant that it does, but this is unproblematic. It’s unproblematic because by far the clearest cases of people having sufficient evidence about truths that we would normally regard as trivial are people with very niche interests. And with such people it
does look like they ought to believe these trivial truths, and would be unjustified if they did not so.

To illustrate, there is a website (www.moviebodycounts.com) which tracks the number of deaths depicted in films. The people who run this website probably do both believe that 836 deaths are depicted in *Lord of the Rings: Return of the King*, and have sufficient evidence that there are (since they list that on their website). The best means for these people to conform to DOXASTIC ATTITUDE with regards to that proposition would, I grant, be to believe it. But this is unproblematic. If they disbelieved that proposition, or suspended judgement about it, they would look unjustified in doing so. Likewise, if trainspotters, stamp collectors, or hobbyists of other varieties failed to believe what their evidence supports about the niche subject matters they are interested in, they would also look unjustified.

Perhaps there are some other cases of trivial truths that are problematic for DOXASTIC ATTITUDE, i.e. cases in which a) one has some doxastic attitude about a trivial truth, p, b) one has sufficient evidence that p, and c) one would not look unjustified in failing to believe that p. But the burden of proof is on the critic of DOXASTIC ATTITUDE to provide such a case.

In summary, in cases of trivial truths which DOXASTIC ATTITUDE prescribes believing, it is either implausible or plausible that one ought to believe them. Where this is implausible (like the truth about the nationality of the bronze medal winner), I have argued that one inevitably doesn’t have sufficient evidence, so that the means one should take to conform to DOXASTIC ATTITUDE will not involve believing them. Where this is plausible, (like the truth about how many deaths are depicted in *Lord of the Rings* for those who run the website), it is also plausible that one ought to believe such truths, even if they are trivial.

In conclusion, although DOXASTIC ATTITUDE prescribes believing some trivial truths, the means one should take to conform to it will never involve believing trivial truths in any way that looks problematic.

5. Conclusion

So should I believe all the truths? In defending DOXASTIC ATTITUDE, I’ve defended a qualified affirmative answer of ‘yes’. One should believe all and only the truths one has some doxastic attitude about. That is the only reformulation of a truth norm that avoids
the implausible aspects of a prescription to believe all and only the truths, but without thereby becoming incapable of explaining justification.

In doing this, I hope I’ve also succeeded in suggesting a new direction for the philosophical literature on how to formulate a truth norm. The literature has often treated the formulation of a truth norm as if it were a logical exercise, successful if the resulting truth norm doesn’t prescribe believing any problematic truths of one kind or another. This cannot be the whole story if a truth norm is going to do any explanatory work, and not be a mere curiosity. We also need to keep in mind whether the truth norm we end up defending can also still explain justification.
This chapter suggests that a truth norm has a particular difficulty explaining when and why one ought to suspend judgement. For a truth norm to succeed in explaining justification, the subsidiary norms which one should follow as a means of conforming to that norm need to prescribe suspending judgement in many cases. However, given that one can only succeed in conforming to a truth norm by having a true belief, it’s difficult to explain why the means of conforming to a truth norm would ever involve suspending judgement. This problem has not been fully appreciated by defenders of truth norms, and the main answers so far given, I’ll suggest, are inadequate. I’ll conclude that a truth norm can in the end explain when and why one ought to suspend judgement, but in order to do so we need to think that there is another fundamental epistemic norm in addition to whatever particular truth norm we opt for. This additional fundamental epistemic norm is one that states that if one fails to conform to this truth norm, one should not do so by having a false belief. In other words, we need to think that there is a contrary-to-duty obligation to err on the side of caution.

1. The suspension-of-judgement problem

Appealing to a truth norm is an attempt at explaining epistemic justification by thinking of being justified as a matter of meeting epistemic norms. Specifically, it claims that being justified is a matter of meeting subsidiary epistemic norms that are a means to the end of conforming to a fundamental truth norm. It is sometimes neglected that this kind of explanation must not only explain when one ought to believe or disbelieve p (i.e. believe that not-\(p\)), but also explain when one ought to suspend judgement about whether p.

If an explanation of justification in terms of a truth norm is to be at all plausible, the subsidiary epistemic norms are going to have to prescribe suspending judgement in many cases. This is because suspension of judgement is part of the story of what it is to be justified. To illustrate, consider whether the pseudonymous author Elena Ferrante is a woman. I’ve suspended judgement about this. Given the absence of evidence about the
author’s gender, it seems that if I were to believe that Elena Ferrante is a woman, that belief would be unjustified. Therefore, if a fundamental truth norm is going to explain justification, then following the subsidiary norms – which I should follow as means to the end of conforming to a truth norm – is going to have to involve suspending judgement in this case, and in many other cases.

A truth norm, however, faces a particular difficulty in explaining when and why one ought to suspend judgement. The key reason is because a truth norm doesn’t prescribe ever suspending judgement. This point was made by Bernard Mayo in response to the following early version of a truth norm put forward by Allen Phillips Griffiths:

It is wrong to believe what is false, and right to believe what is true. Whatever else one does with a truth, believing the proposition that expresses it is the first and most fitting thing to do with it. (Phillips Griffiths 1962, 182)

Mayo points out that Phillips Griffiths’ claim has the consequence that one ought to never suspend judgement:

An immediate consequence of the thesis that believing is what is fitting to do with a truth is, of course, that one ought to believe the truth. Further, one ought never to suspend judgement, since there is nothing that it is right not to believe, except the false, the negation of which, being true, one ought to believe. (Mayo 1963, 144)

The difficulty created by the fact that a truth norm prescribes never suspending judgement has not been fully appreciated by proponents of truth norms. We can illustrate this difficulty in relation to one of the simpler truth norms we’ve discussed,

---

1 This was written before Elena Ferrante was ‘outed’ by a journalist. Perhaps now it’s no longer the case that one should suspend judgement about this proposition, but my use of this example should be understood relative to the evidential situation before this took place. Relative to that evidential situation, I take it that it is plausible that one ought to suspend judgement about whether Elena Ferrante is a woman.

2 The question of how a truth norm accounts for suspension of judgement hasn’t received much discussion. It is briefly discussed by Wedgwood (2002, 272–73; cf. 2008, 9; 2013a, 232; 2013b, 126) and Engel (2013a, 213). However, as I’ll discuss, neither of them recognise the full extent of the problem suspension of judgement creates for a truth norm. I think Conor McHugh is the person writing in the truth norm literature who has come closest to understanding the real problem that suspension of judgement raises (2012, 15–16). I’ll discuss McHugh further in fn. 2 below.
Although I rejected OUGHT in favour of DOXASTIC ATTITUDE, in this chapter, it will simplify things to discuss the suspension-of-judgement problem in relation to OUGHT. Because of this, I’m going to discuss possible solutions to the suspension-of-judgment problem, and argue for my own solution, primarily in relation to OUGHT. However, everything I say in this chapter could be said mutatis mutandis in relation to DOXASTIC ATTITUDE. And from the point of this thesis in its entirety, I would want to defend a version of this chapter’s conclusion that is relative to DOXASTIC ATTITUDE.

So how does suspension of judgement raise a problem for OUGHT? Like any truth norm, OUGHT is supposed to explain justification because one should follow subsidiary norms as a means to the end of conforming to it. Suspending judgement looks strange in this context, because it’s hard to see why one should ever suspend judgement as a means to end of conforming to OUGHT. This is because OUGHT is a norm that I can only succeed in conforming to by having a true belief about p; this is Mayo’s point made in reference to OUGHT.

Suspending judgement about whether p, however, is something which makes it rationally impossible for me to have a true belief about whether p. I say ‘rationally impossible’ because I don’t want to deny that it’s possible – i.e. psychologically possible – in some strange cases for someone to count as both believing that p and suspending judgement about p. An example of such a case would be someone who unconsciously believes that their son has died on the battlefront, but consciously suspends judgement about whether he has. I just want to claim that one is necessarily irrational in such a case. In doing so, I’m making the uncontroversial assumption that if one ought to suspend judgement about whether p (in the subjective sense of ‘ought’ of the subsidiary norms), then it’s not the case that one also ought to believe that p (again in the subjective sense). This is uncontroversial because it is just to say that one cannot at the same time have a justified belief that p and be justified in suspending judgement about whether p.

But given that suspending judgement about whether p is something which makes it rationally impossible for me to have a true belief about whether p, then it’s also something that makes it rationally impossible for me to succeed in conforming to OUGHT. But, as we outlined above, if a truth norm like OUGHT is going to succeed in
explaining justification, then the means one should take to conform to OUGHT must involve suspending judgement in some cases. But how can the means one should take to conform to OUGHT involves doing something that makes it rationally impossible for me to succeed in conforming to OUGHT? In more general terms, if the aim is to believe something if and only if it’s true, how can one be doing what one ought with respect to that aim by doing something – suspending judgement – that makes it rationally impossible that one will achieve the aim?

This is what I’ll call ‘the suspension-of-judgement problem’. An answer to this problem needs to explain why one ought to suspend judgement as means to the end of conforming to a truth norm.

However, because a truth norm is supposed to explain justification – and the suspension-of-judgement problem is a problem with that explanation – not just any explanation of why one ought to suspend judgement will do. This is an important point, and we can illustrate this by describing how opting for a permissive truth norm, such as the following, is not a good answer to the suspension-of-judgement problem:

MAY: One may believe that p if and only if p is true.

MAY does not face the same problem with suspension of judgement as OUGHT. This is because suspending judgement about whether p is not something that makes it rationally impossible to succeed in conforming to MAY, and so there’s no puzzle about why one should suspend judgement in order to conform to MAY.

However, MAY does not thereby provide a good answer to the suspension-of-judgement problem. This is because, as we demonstrated in the last chapter, the best means of conforming to a permissive truth norm is to suspend judgement about everything. This means in effect that the subsidiary norms one should follow as a means of conforming to MAY will prescribe suspending judgement in far too many cases. So although MAY can explain why one ought to suspend judgement, it does not do so in a way that vindicates its explanation of justification. Vindicating the truth norm account of justification doesn’t just involve explaining why one ought to suspend judgement, but explaining why one ought to suspend judgement in the right cases, i.e. in cases in which one would be unjustified either in believing that p or in disbelieving that p.

For the rest of this chapter, I’ll discuss different ways to deal with the suspension-of-judgement problem. Firstly, I’ll consider two attempts from two
proponents of truth norms, Pascal Engel and Ralph Wedgwood. I’ll then put forward my own way of dealing with this problem.

In §2, I’ll discuss Engel’s suggestion that a truth norm can explain why one ought to suspend judgement by appealing to evidential norms which one should follow as a means to the end of conforming with a truth norm. Engel’s answer is insufficient, I’ll argue, because he fails to show that in order to conform to a truth norm, one should follow an evidential norm that prescribes suspending judgement in the right cases. In particular, he fails to show that one should follow a norm that requires one to have what we normally think of ‘sufficient evidence’. Doing this, as I’ll show, involves giving different weights to the two aims of believing the truth and avoiding error; in particular it involves giving greater weight to the aim of avoiding error.

Therefore, in §3, I’ll discuss ways the truth norm account of justification can give the aim of avoiding error greater weight. Firstly, I’ll discuss Wedgwood’s claim that suspending judgement has an ‘intermediate degree of correctness’. This claim, I’ll show, has the result that the aim of avoiding error has greater weight. It might not be entirely obvious why it has this result, and it’s not clear that Wedgwood intends it to, but it does. However, I won’t in the end favour this way of giving the aim of avoiding error greater weight, because I am sceptical of Wedgwood’s appeal to degrees of correctness for belief.

I’ll end by putting forward a different way of giving the aim of avoiding error greater weight. I’ll argue we should think that the fundamental epistemic norms include, in addition to whatever particular truth norm we opt for, an additional contrary-to-duty obligation to err on the side of caution, i.e. an additional norm which states that if one fails to conform to this truth norm, one should not do so by having a false belief. This provides us with a set of fundamental epistemic norms that can explain a higher-threshold evidential norm, and thereby deal with the suspension-of-judgement problem.

2. Engel’s appeal to evidential norms

Engel suggests that a truth norm can explain why one ought to suspend judgement by appealing to evidential norms. Evidential norms, Engel suggests, are norms which both prescribe suspending judgement in the right cases and which one should follow as a means of conforming to the fundamental truth norm.

Engel claims, like many other proponents of truth norms (see e.g. Boghossian 2003, 39; Shah 2006; Whiting 2010, 221–22; McHugh 2014b), that if one is trying to
conform to a truth norm, one should follow evidential norms that prescribe basing one’s beliefs on evidence:

There are two candidates for the correctness condition of belief. One is the external, or as it is sometimes called, objective norm of truth. The other is the internal norm, sometimes called the subjective norm, which associates correct belief with our evidential reasons for believing. They seem to be closely related, for it seems that one cannot follow the first without following the second: the best way to know the truth is to rely on one’s evidence, and evidence is evidence for truth. (Engel 2013a, 207)

And this means, Engel claims, that a truth norm can explain why one ought to suspend judgement. This is because following evidential norms – which one should follow as a means to the end of conforming to a truth norm – involves suspending judgement in some cases. In particular, they prescribe suspending judgement in cases in which one doesn’t have enough evidence for either p or not-p:

It is false that the norm of truth allows only two doxastic attitudes. If one considers whether p is true, and does not have enough evidence for either p or not-p, the norm does not prescribe believing p or believing not-p. It prescribes withholding belief. But isn’t withholding belief then under the governance of the evidential norm or the norm of i-correctness of justification? It is, but it is also under the governance of the truth norm, for there is no possibility of being governed by the truth norm unless one follows the evidential norm. (Engel 2013a, 213)

I don’t think Engel’s response gets to the heart of the matter. I agree with Engel that belief is subject to evidential norms that prescribe believing that p only if one has enough or sufficient evidence. The following evidential norm, which I introduced in Chapter 3, is an example:

SUFFICIENT EVIDENCE: One ought to believe that p if and only if one has sufficient evidence that p.

And I admit that if Engel is right that one should follow this kind of evidential norm in order to conform to a truth norm like OUGHT, that would deal with the suspension-of-judgement problem. This is because SUFFICIENT EVIDENCE, unlike a permissive truth norm, prescribes suspending judgement in the right cases, because it prescribes suspending judgement in cases in which one doesn’t have sufficient evidence either in support of p or in support of not-p.
However, I don’t think Engel has given good reason to think that one should follow an evidential norm like SUFFICIENT EVIDENCE in order to conform to OUGHT. In fact, I think it’s question-begging to claim, as Engel does, that one should use an evidential norm like SUFFICIENT EVIDENCE in order to conform to OUGHT. This is because the same fact that made suspension of judgement a problem in the first place – the fact that it’s rationally impossible to conform to a truth norm by suspending judgement – also gives us reason to doubt the claim that one should follow an evidential norm like SUFFICIENT EVIDENCE as a means of conforming to OUGHT. In other words, if the aim is to believe something if and only if it’s true, why should one do something – following SUFFICIENT EVIDENCE – which makes it rationally impossible to achieve that aim in cases in which one doesn’t have sufficient evidence either for p or for not-p?

Therefore, in the context of answering the suspension-of-judgement problem, one cannot assume, as Engel does, that the means one should take to conform to a truth norm like OUGHT involves following a norm like SUFFICIENT EVIDENCE, because the suspension-of-judgement problem should equally make us doubt that.

I'll demonstrate this more fully in two steps. Firstly, in §2.1, I'll go into more detail about what it is to have ‘sufficient evidence’ and consider other alternative evidential norms one could follow. In particular, I'll contrast SUFFICIENT EVIDENCE with the following evidential norm:

MORE THAN 50-50: One ought to believe that p if and only if one’s evidence makes p more likely than not-p.

MORE THAN 50-50 prescribes believing that p when one has a much lower threshold of evidence in support of p than what SUFFICIENT EVIDENCE requires.

Secondly, I'll argue in §2.2 that a truth norm like OUGHT can only succeed in explaining MORE THAN 50-50 and not SUFFICIENT EVIDENCE. Although I'll make this point with reference to OUGHT in particular, the point generalises. OUGHT can only explain MORE THAN 50-50, I'll argue, because of the way it balances the dual aims of believing the truth and avoiding error. OUGHT cannot, I'll argue, justify giving anything other than equal weight to these two aims. And this means that following MORE THAN 50-50 is better than following SUFFICIENT EVIDENCE as a means of conforming to OUGHT.
MORE THAN 50-50, however, is not an evidential norm that prescribes suspending judgement in the right cases, because following it would result in beliefs that look unjustified. Therefore, if I’m right that MORE THAN 50-50 is better than SUFFICIENT EVIDENCE as a means of conforming to OUGHT, then merely appealing to evidential norms does not answer the suspension-of-judgement problem. Moreover, the reason why MORE THAN 50-50 is better than SUFFICIENT EVIDENCE as a means of conforming to OUGHT is, as I’ll show, essentially the same reason why suspension of judgement created a problem in the first place.

2.1. What it is to have ‘sufficient’ evidence

What does having sufficient evidence involve? It is, as we briefly sketched in the last chapter, supposed to express an intuitive requirement on justification. In order to have a justified belief that p, it is not enough for one to have some evidence that supports p. For example, that Clifford is barking at the door provides me with some evidence that the postman is at the door. But it’s not enough for me to have a justified belief that the postman is at the door, because Clifford could also be barking because my brother has arrived home, or because someone is walking on the pavement outside the house. Having a justified belief that the postman is at the door requires me to have sufficient evidence that he is, which in this case probably will just involve me looking out the window. This is the requirement that is encoded in SUFFICIENT EVIDENCE.

To go into a bit more detail, it seems that to have sufficient evidence that p one’s evidence needs to support p above certain threshold of strength. How strongly does my evidence need to support p in order to have sufficient evidence that p? Although I don’t want to give an unrealistically precise answer to this question, we can say that having sufficient evidence falls between two extremes.

We can illustrate these extremes in terms of probability. I’m going to assume that how strongly a body of evidence E supports p is how likely p is given E. One may object to identifying evidential support with probability (see e.g. Achinstein 2001, chaps. 4–6). But any account of evidence will need some way of quantifying how strongly evidence supports a hypothesis. Therefore, we’ll still be able to characterize having sufficient evidence in much the same way as I do below mutatis mutandis for any alternative way of quantifying evidential support. In any case, it looks like we’re going to have to identify evidential support with probability if we are suggesting that a truth norm explains why
belief is subject to evidential norms. This is because this explanation is going to have to involve the claim that one should base one’s beliefs on evidence because doing so increases the probability of believing truly. This seems to be what is suggested by Engel’s remark that ‘evidence is evidence for truth’.

So if we assume that a body of evidence E supports p to the degree that it makes p likely, how likely does p need to be given my evidence in order for me to have sufficient evidence? The answer seems to be that sufficient evidence supports p above a certain threshold of likelihood, a threshold somewhere in between the two extremes of 50% and 100% likelihood.

Having sufficient evidence that p at the very least requires that p is more likely than not – i.e. more than 50% likely – given my evidence. But that by itself doesn’t look enough for sufficient evidence. It doesn’t look like I have sufficient evidence that p – i.e. evidence that would make me justified in believing that p – if p is just a bit more than 50% likely given my evidence.

We can illustrate this with the following example given by Michael Pace:

[S]uppose that I have an unfair coin that I know, statistically, lands heads 55% of the time. Even though such evidence makes it somewhat more likely than not that the coin will land heads on one particular toss, the evidence still seems insufficient for me to be epistemically justified in forming an outright belief that the coin will land heads on an individual toss.

(Pace 2011, 246)

Although for me to have sufficient evidence that p requires that p is more likely than not given my evidence, this case illustrates that this is not enough for me to have sufficient evidence that p. In fact, having sufficient evidence that p would normally seem to require p to be a fair bit more than 50% likely given my evidence.

On the other hand, having sufficient evidence that p does not look like it requires p to be absolutely certain given my evidence, i.e. 100% likely given my evidence. To use Dretske’s example, it seems that I can have sufficient evidence that, say, an animal is a zebra even if my evidence doesn’t make it 100% likely that it’s a zebra, because my evidence doesn’t rule out with absolute certainty the hypothesis that it’s a cleverly disguised mule (1970, 1016).

Therefore, having sufficient evidence that p does not require that p is 100% likely given my evidence. In fact, it seems that one still can have sufficient evidence that p if one’s evidence supports p to a degree a fair bit lower than 100% certainty.
These two cases together illustrate that sufficient evidence is evidence that makes p likely to a degree that is in between these two extremes. In other words, the likelihood of p given my evidence needs to be above a certain threshold – a threshold in between 50% and 100% - in order for me to have sufficient evidence that p.

Apart from saying this, I’m not going to define this threshold – i.e. give an exact figure of how likely p needs to be given one’s evidence. But even if we can’t precisely define this threshold, it still seems that it is an intuitive requirement on having a justified belief that one’s evidence supports p above some threshold of likelihood. This is, in brief, the requirement encoded in SUFFICIENT EVIDENCE.

However, SUFFICIENT EVIDENCE is not the only possible evidential norm. Compare SUFFICIENT EVIDENCE with the alternative evidential norm I introduced earlier:

MORE THAN 50-50: One ought to believe that p if and only if one’s evidence makes p more likely than not-p.

MORE THAN 50-50 is still an evidential norm. It still recommends basing one’s beliefs on and only on the evidence. The key difference between it and SUFFICIENT EVIDENCE is that MORE THAN 50-50 requires a much lower threshold of evidential support than SUFFICIENT EVIDENCE.

This difference between these two evidential norms is illustrated by the following two tables:

<table>
<thead>
<tr>
<th>Probability</th>
<th>SUFFICIENT EVIDENCE</th>
<th>MORE THAN 50-50</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Believe</td>
<td>1</td>
</tr>
<tr>
<td>0.5</td>
<td>Suspend judgement</td>
<td>0.5</td>
</tr>
<tr>
<td>0</td>
<td>Disbelieve</td>
<td>0</td>
</tr>
</tbody>
</table>

158
The reason why I’m comparing these two evidential norms is that only one of them, SUFFICIENT EVIDENCE, can provide a good answer to the suspension of judgement problem. This is because only SUFFICIENT EVIDENCE prescribes suspending judgement in the right kinds of case, i.e. cases in which one would be unjustified in either believing or disbelieving p.

MORE THAN 50-50, on the other hand does not prescribe suspending judgement in the right cases. This is because MORE THAN 50-50 prescribes believing that p if one’s evidence makes p more likely than not, and disbelieving p – i.e. believing that not-p – if one’s evidence makes p less likely than not. Because of this, MORE THAN 50-50 only prescribes suspending judgement about whether p in cases, like fair coin flips, in which p is exactly 50% likely given my evidence. However, as I said before, someone who believes that p when p is only slightly more likely than not given their evidence intuitively looks unjustified. And that means that MORE THAN 50-50, because it prescribes belief rather than suspending judgement in such a case, does not prescribe suspending judgement in the right cases.

Therefore, in order for Engel’s appeal to evidential norms to provide a satisfactory answer to the suspension-of-judgement problem – i.e. one that vindicates an explanation of justification in terms of a truth norm – SUFFICIENT EVIDENCE needs to be better than MORE THAN 50-50 as a means of conforming to OUGHT. However, as I’ll show in the next section, it is not.

2.2. Why following MORE THAN 50-50 is a better means of conforming to OUGHT

In this section I’ll show that following a lower-threshold evidential norm like MORE THAN 50-50 is in fact a better means of conforming to a truth norm than a higher-threshold evidential norm like SUFFICIENT EVIDENCE. Again, I’ll make this point in reference to OUGHT, but I’ll outline afterwards how the point generalises.

Because OUGHT is a biconditional truth norm – i.e. one that prescribes believing that p if and only if p is true – someone trying to conform to it is trying to balance two competing aims. These dual truth aims are the “two ways of looking at our duty in the matter of opinion” discussed by William James: “Believe truth! Shun error!” (1897, 17–18).

The fact that trying to conform to OUGHT involves balancing the dual aims of believing the truth and avoiding error does mean, I’ll argue, one should have some evidence
threshold – i.e. it means that one should believe that \( p \) if and only if one’s evidence supports \( p \) above some minimum threshold of strength – as a means to the end of conforming to \textit{ought}.

However, having the lower evidence-threshold of \textit{more than} 50-50 – and not the higher one of \textit{sufficient evidence} – is the better means of conforming to \textit{ought}. The key reason behind this second claim is that \textit{ought} can only justify giving \textit{equal weight} to the dual aims of believing the truth and avoiding error. And this means that \textit{more than} 50-50 is the better means of conforming to \textit{ought}. \textit{Sufficient evidence} would be the better means if we gave \textit{greater} weight to the aim of avoiding error, but \textit{ought} cannot, by itself, justify this.

I’ll start by explaining why trying to conform to \textit{ought} involves having some evidence threshold. This is because, as I’ve said, \textit{ought} involves the dual aims of believing the truth and avoiding error.

That \textit{ought} involves these dual aims can be made apparent if we recall its phrasing:

\textit{ought}: One ought to believe that \( p \) if and only if \( p \) is true.

\textit{Ought} is the conjunction of two separate norms. Firstly, the following norm:

\textit{if}: One ought to believe that \( p \) if \( p \) is true.

And secondly, the following norm, which was one of the permissive truth norms we discussed in the last chapter:

\textit{only if}: One ought to believe that \( p \) only if \( p \) is true.

Someone trying to conform \textit{ought} is not only trying to ensure that they have true beliefs – i.e. to conform to \textit{if}. They are also trying to ensure that they avoid having false beliefs – i.e. to conform to \textit{only if}. Trying to conform to \textit{ought} involves balancing \textit{both} of these two aims.

This means that with regards to any proposition, there are two distinct ways of \textit{failing} to conform to \textit{ought}. Firstly, one can fail by failing to have a true belief, which is a failure according to \textit{if}. Secondly, one can fail by having a false belief, which is a failure according to \textit{only if}.

The fact that \textit{ought} involves balancing these two aims means one should
follow an evidential norm that prescribes believing that p if and only if one has evidence over some threshold. This can be shown by thinking about what would be the case if we only had one of the two aims and not the other. In either case, it would not make sense to have any evidence threshold.

On the one hand, say I only have the aim of believing truly, i.e. I am only trying to conform to IF, and not to ONLY IF. Then it would not make sense for me to have an evidence threshold. Rather, if I only had the aim of believing truly, then, with regards to any proposition p, I should both believe that p and believe that not-p, because that guarantees that I'll succeed in the aim of believing truly. This strategy would, of course, also result in me also having a false belief about whether p. But that doesn’t matter if I only have the aim of believing truly and not the aim of avoiding error. Therefore, it wouldn’t make sense to have any evidence threshold if I only had the aim of believing truly, because any threshold would be too high.

On the other hand, say I only have the aim of avoiding believing falsely, i.e. I’m only trying to conform to ONLY IF, and not to IF. If I’m doing this, it again doesn’t make sense for me to have an evidence threshold. Rather, if I only had the aim of avoiding believing falsehoods, then, as we saw in the last chapter, I should suspend judgement about everything. This would mean that I wouldn’t end up with any true beliefs. But this does not matter if my only aim is ensuring that I don’t have false beliefs. Suspending judgement is just as much a way of avoiding a false belief as having a true belief. Therefore, if I were just trying to avoid error, then it would not make sense for me to have any evidence threshold, because no threshold would be high enough.

But if I am trying to conform to OUGHT – i.e. if I’m trying to both ensure that I believe truths and ensure that I avoid error – then the seemingly negative aspects of these two strategies do matter. Believing p and believing not-p is not a good way of balancing these two aims and nor is suspending judgement about everything. Therefore, if I am trying to balance these two aims, then it would make sense for me to have an evidence threshold. To balance these two aims, I should follow an evidential norm that prescribes believing p if my evidence supports p above a certain degree of strength – to give me a good chance of believing some truths – but only if my evidence supports above that degree of strength – to give me a good chance of not believing falsehoods. Therefore, trying to conform to OUGHT will involve having some evidence threshold.

However, as I’ll now show, the lower evidence threshold of MORE THAN 50-50 is
better than SUFFICIENT EVIDENCE as a means of conforming to OUGHT.

The key reason behind this latter claim is that in order to justify the higher evidence threshold of SUFFICIENT EVIDENCE, we would need to grant greater weight to the aim of avoiding error than to the aim of believing the truth. And OUGHT cannot, by itself, justify granting greater weight to either of the dual Jamesian aims. To give greater weight to one of two two aims would be to claim that one way of failing to conform to OUGHT – i.e. having a false belief or failing to have a true belief – is preferable to the other. Given that OUGHT is a deontic norm, which one either succeeds or fails to conform to as a binary matter, it doesn’t look like it’s a norm that can justify the claim that one of these two ways of failing to conform is preferable to the other. Therefore, if OUGHT is the sole fundamental norm on belief, we have to assume that the two aims of believing the truth and avoiding error have equal weight.3

But if that’s the case, then MORE THAN 50-50 is better than SUFFICIENT EVIDENCE as a means of conforming to OUGHT. I’ll illustrate with the proposition that Elena Ferrante is a woman. If we assume that there is equal weight attached to the aim of having a true belief about whether Elena Ferrante is a woman to the aim of avoiding a false belief about whether Elena Ferrante is a woman, then following MORE THAN 50-50 with regards to this proposition looks like the better strategy to balance these two aims than SUFFICIENT EVIDENCE. As we saw earlier, if one only gave weight to the aim of having a true belief, then one should both believe that Elena Ferrante is a woman and believe that that’s not the case, because that guarantees true belief. And if one only gave weight to the aim of avoiding error, then one should only ever suspend judgement about whether Elena Ferrante is a woman, because that guarantees not having a false belief.

Given these two extremes, if one gave these two aims equal weight, following MORE THAN 50-50 – and not SUFFICIENT EVIDENCE – looks like the happy medium. If

3 McHugh makes a point very much like this when he claims that a truth norm like OUGHT entails that belief is the ‘uniquely correct doxastic attitude’ to have towards a truth (2012, 15). This means, McHugh claims, that suspending judgement about a truth, and disbelieving it, are equally incorrect (2012, 16). This is in essence the same as my claim that OUGHT cannot give the aim of avoiding error greater weight. McHugh does not discuss, however, implications this has for evidential norms. But the problem McHugh raises does mean that a criticism I have made elsewhere of McHugh’s discussion of suspension of judgement – in which I argued that merely appealing to the dual Jamesian aims suffices to explain suspension of judgement (Greenberg and Cowie 2016, 13–14) – was overly simplistic.
having a true belief about whether Elena Ferrante is a woman is equally important as avoiding a false one, then it seems that the best way of striking a balance between these two aims is to believe that Elena Ferrante is a woman if and only if it's more likely than not.

We can give two main reasons to back this up. Firstly, it’s clear that following MORE THAN 50-50 is not a strategy that’s worse than following SUFFICIENT EVIDENCE from the point of view of balancing these dual aims. Compared with SUFFICIENT EVIDENCE, following MORE THAN 50-50 admittedly has the drawback that it makes it more likely that one will have a false belief about whether Elena Ferrante is a woman – because it requires a lower threshold of evidential support. But for the same reason following MORE THAN 50-50 also has the advantage that it makes it more likely that one will have a true belief about whether Elena Ferrante is a woman. And it doesn’t seem that the increase in the likelihood of believing falsely from following MORE THAN 50-50 could be greater than the correlative increase in the likelihood of believing truly. And if the aims of believing the truth and avoiding error have equal weight, then any ill effect following MORE THAN 50-50 has with regards to the aim of avoiding error is ‘paid for’ equally by the beneficial effect it has with regards to the aim of believing the truth.

Secondly, and more importantly, following SUFFICIENT EVIDENCE has a crucial disadvantage when compared to MORE THAN 50-50. This is because following SUFFICIENT EVIDENCE involves suspending judgement in many more cases than following MORE THAN 50-50. Whereas following MORE THAN 50-50 involves suspending judgement about whether p only in fairly exceptional cases – those in which p is exactly 50% likely given my evidence – following SUFFICIENT EVIDENCE involves suspending judgement in any case in which one doesn’t have sufficient evidence either for p or for not-p. This means that following SUFFICIENT EVIDENCE involves suspending judgement in a far greater number of cases.

Because of this, in many more cases than with MORE THAN 50-50, following SUFFICIENT EVIDENCE involves doing something – suspending judgement – which guarantees it that one won’t achieve the aim of believing the truth. On the other hand, in cases in which one’s evidence does not support p above the threshold for sufficiency, but supports p more than its negation, following MORE THAN 50-50 at least gives one the possibility of achieving the aim of believing the truth. Therefore, because following SUFFICIENT EVIDENCE involves ruling oneself out of achieving the aim of believing the truth in a way that following MORE THAN 50-50 does not, we can conclude that MORE THAN
50-50 is the evidential norm one should follow if the dual aims of believing the truth and avoiding error have equal weight.

This conclusion undermines the answer Engel gave to the suspension-of-judgement problem. It does so because it means that Engel’s appeal to evidential norms does not provide an answer to the suspension-of-judgement problem that vindicates the truth norm account of justification. This is because more than 50-50 is not an evidential norm which prescribes suspending judgement in the right cases, for in cases in which the evidence makes p only slightly more likely than its negation, it doesn’t look like one is justified in believing that p.

We can illustrate again with the example of Elena Ferrante. It does seem that the evidence makes it more likely that Elena Ferrante is a woman than not, e.g. it’s more likely, statistically speaking, for a woman to have the insight that Elena Ferrante’s novels manifest of what it’s like for a girl growing up. However, that doesn’t seem to mean that I would be justified in believing that Elena Ferrante is a woman. Therefore, if the subsidiary evidential norm one should follow to conform to a truth norm is more than 50-50, then that truth norm is not going to succeed in explaining justification.

Furthermore, the reason why more than 50-50 is a better means than sufficient evidence of conforming to ought is essentially the same reason why suspension of judgement created a problem in the first place. Suspension of judgement created a problem because if the aim is to believe something if and only if it’s true, then it looks strange why one should do something – suspending judgement – which makes it rationally impossible for one to achieve that aim. More than 50-50 is a better means than sufficient evidence of conforming to ought for essentially the same reason, because if the aim is to believe something if and only if it’s true, it then looks strange as to why one would try to do that by doing something – following sufficient evidence – which, in cases in which one doesn’t have sufficient evidence either way, makes it rationally impossible to achieve that aim. This is especially strange if there’s an alternative – following more than 50-50 – that does not result in it being rationally impossible to achieve that aim in the same cases.

In this way, more than 50-50 is a better means of conforming to ought for essentially the same reason why suspension of judgement created a problem in the first place. Therefore, it looks like Engel has left the fundamental problem created by suspension of judgment unanswered.
3. The solution: give greater weight to the aim of avoiding error

We’ve seen that an appeal to evidential norms does not by itself answer the suspension-of-judgement problem. But if a truth norm could justify an evidential norm like SUFFICIENT EVIDENCE, then, I admit, the suspension-of-judgement problem could be answered. Whereas conformity with MORE THAN 50-50 does not make for justification, it is plausible that conformity with SUFFICIENT EVIDENCE does make for justification. And the key difference is that following SUFFICIENT EVIDENCE involves suspending judgement in the right kinds of cases, whereas following MORE THAN 50-50 did not.

The key reason why SUFFICIENT EVIDENCE could not be explained by OUGHT was because that would require giving greater weight to the aim of avoiding error. OUGHT could not do this, because it could only justify giving equal weight to the dual Jamesian aims.

This opens up a simple solution to the suspension-of-judgement problem: we should find some way to give greater weight to the aim of avoiding error than to the aim of believing the truth. If we could do that, then we could make sense of why one should follow an evidential norm like SUFFICIENT EVIDENCE, the following of which involves suspending judgement in the right kind of cases.

In this rest of this chapter, I’m going to discuss two different ways of granting greater weight to the aim of avoiding error. Firstly, in §3.1 I’ll discuss a claim, made by Ralph Wedgwood, that suspension of judgement has an ‘intermediate degree of correctness’ in between the complete correctness of true belief and the complete incorrectness of false belief. Although it’s not clear that he intends it, in making this claim Wedgwood in effect gives the aim of avoiding error greater weight. In this respect, Wedgwood’s account is on the right track. In the end, however, I’ll prefer a different way, because of problems with Wedgwood’s appeal to degrees of correctness.

In §3.2, I’ll put forward a different way of giving the aim of avoiding error greater weight. I’ll suggest we do this by holding that there is, in addition to whatever truth norm we want to defend, a contrary-to-duty obligation to err on the side of caution, i.e. an obligation that states that if one fails to conform to that truth norm, one shouldn’t do so by having a false belief. This additional norm, I’ll argue, enables us to explain why there is an evidential norm with a higher threshold of evidential support – like SUFFICIENT EVIDENCE – and thereby answer the suspension-of-judgement problem.
3.1. Wedgwood: an intermediate degree of correctness

Wedgwood’s claim that suspending judgement has an intermediate degree of correctness comes in the context of the particular truth norm Wedgwood defends, which is stated in terms of ‘correctness’, rather than the deontic vocabulary of ‘ought’ that I have been using:

CORRECT: A belief that p is correct if and only if p is true. (Wedgwood 2002, 268)

Wedgwood claims that this norm explains justification in the same kind of way that we have discussed in detail in relation to other truth norms. And why one ought to suspend judgement, according to Wedgwood’s account, is because suspending judgement has an intermediate degree of correctness in between the complete correctness of true belief and the complete correctness of incorrectness of false belief:

According to the fundamental norm of correct belief, I propose, suspending judgment about p is neither correct nor incorrect. If one suspends judgment about p then one has neither got things right nor got things wrong about p. Thus, this fundamental norm ascribes to the state of suspending judgment about p an intermediate value, between the complete correctness of believing p when p is true, and the complete incorrectness of believing p when p is not true. (Wedgwood 2002, 272–73; cf. 2008, 9; 2013a, 232; 2013b, 126)

To claim that suspension of judgement has an intermediate degree of correctness is in effect to give the aim of avoiding error greater weight than the aim of believing the truth. Initially, it might not be entirely clear why Wedgwood’s claim has this consequence. It does so because if suspending judgement is more correct than having a false belief, then one way of failing to conform to the truth norm – not having a true belief – is preferable to the other – having a false belief. And as we saw earlier, this can only be the case if the aim of avoiding error trumps the aim of believing the truth. Therefore, claiming that suspension of judgement has an intermediate degree of correctness, though it might not be obvious, entails that the aim of avoiding error has greater weight than the aim of believing the truth.

Although Wedgwood commits himself in this way to avoiding error having greater weight, he doesn’t say this explicitly. And the important point is that it has to be thought of as an addition to the claim that belief is subject to a truth norm. The idea that suspension of judgement has an intermediate degree of correctness is not contained in
the content of the truth norm Wedgwood defends, CORRECT.

The fact that it’s an addition shouldn’t put us off making this claim. We have reason to give the aim of avoiding error greater weight because otherwise we can’t explain suspension of judgment within the truth norm account of justification. But we just need to be conscious that it goes beyond the claim that belief is subject to a truth norm.

Furthermore, I’m not sure Wedgwood intends his claim that suspending judgement has an intermediate degree of correctness to result in the aim of avoiding error having greater weight, because it runs contrary to things he says elsewhere. In particular, Wedgwood also claims that we also need to appeal to practical considerations to fully explain how much one’s evidence needs to support to support p for one to count as justified in believing that p. Wedgwood claims we need to appeal to practical considerations here because of a point he makes about the dual competing truth aims we have already discussed. The fact that the truth aim in fact involves these dual competing goals means, Wedgwood claims, that there can be no purely epistemic answer about how to balance them:

This “goal” in fact involves at least two goals – the goal of having an outright belief in p if p is true, and the goal of not having an outright belief in p if p is false. These goals are distinct: if p is true, then one can achieve the second goal, but not the first, by suspending judgment and neither believing nor disbelieving p at all. But there does not seem to be any purely epistemic principle that can tell us how to balance these two goals against one another.

(Wedgwood 2008, 5)

This claim that there is no purely epistemic answer about how to balance the dual Jamesian aims runs contrary to Wedgwood’s earlier claim that suspending judgement has an intermediate degree of correctness. Wedgwood seems to regard the claim about an intermediate degree of correctness as a purely epistemic claim, as indicated by him saying, in the first passage I quoted, that ‘the fundamental norm ascribes to the state of suspending judgement an intermediate correctness value’. And since we’ve shown that suspending judgement having an intermediate degree of correctness results in the aim of avoiding error having greater weight, that means that, contrary to what Wedgwood says, there is a purely epistemic answer about how to balance these two aims, i.e. avoiding error trumps believing the truth.

Therefore, although Wedgwood does in effect give the aim of avoiding error
greater weight, he perhaps does so unintentionally. And although I think we need to give
the aim of avoiding error to solve the suspension of judgement problem, I’m not going
to do it in the way Wedgwood does. The main reason is because appealing to degrees of
correctness, as Wedgwood does, is problematic.

An initial reason why it’s problematic, given by Engel, is that correctness doesn’t
seem to be a graded concept: “correctness, as we ordinarily use this notion, is not a
matter of degree” (Engel 2013a, 210). This claim of Engel’s seems plausible. The use of a
term like ‘correct’ that figures in Wedgwood’s truth norm is – as an objective normative
concept – most familiar with regards to things like ‘correct’ answers to a quiz or exam. In
these cases we don’t tend to think of answers as being more or less correct or correct to
a certain degree; they are either correct or incorrect. And we definitely don’t think of not
giving an answer – presumably the equivalent of suspending judgement – as being more
correct than a wrong answer and less correct than a right answer. In this way, Engel’s
claim that correctness is not a graded concept looks plausible.

Wedgwood responds to Engel’s criticism by claiming that his account can be
made consistent with the phenomena Engel points to:

Engel points out, in effect, that in ordinary language, adjectives like ‘right’ and ‘correct’ do
not have comparative forms. This is true, but it is compatible with the suggestion that to call
something ‘correct’ is normally to say that it is perfectly or completely correct – that is, that its
degree of incorrectness is 0 (it is not incorrect to any degree at all). (Wedgwood 2013a, 224)

While I don’t think Engel has a knockdown argument against Wedgwood’s appeal to
degrees of correctness, I don’t think that Wedgwood has really got to the heart of the
problem.

The problem Engel is raising is not, I take it, that Wedgwood’s account cannot
possibly be made consistent with ordinary language about correctness, but that once one
claims that there are degrees of correctness, one loses the intuitive basis one had for the
account. In other words, admitting that we don’t normally speak of things as being
correct to a certain degree has the consequence that even if it’s plausible that true beliefs
are correct and false beliefs are incorrect, it cannot be equally plausible that suspension
of judgement has an intermediate degree of correctness. And I don’t think Wedgwood
has provided a satisfactory answer to this.

Where I do think it makes some sense to speak of degrees of correctness in
relation to truth is in talk of credences. If p is true, I think it makes sense to think of a
0.7 credence in \( p \) as more correct than a credence of 0.5, because we can think of it being ‘closer to the truth’. And some formal epistemologists, following James M. Joyce (1998), think we should understand the rational norms on credences (e.g. norms of probabilistic coherence, the Principal Principle, Bayes’ rule) as explained by a norm that one should maximize the closeness to truth of one’s credences (see e.g. Pettigrew 2016). However, such an approach can only be applied to all-out doxastic states (belief, disbelief, and suspension of judgement) if something like the threshold view of belief is correct, i.e. if all-out belief is credence above a certain threshold. And, as I’ve already argued in Chapter 3, the threshold view of belief is implausible, and we are better off not appealing to it. Therefore, it wouldn’t help Wedgwood at this point to appeal to credences (as he sometimes does, e.g. see 2013a, secs. IV–IV) to make the case that suspension of judgement has an intermediate degrees of correctness.

More generally, I think this dispute between Wedgwood and Engel should make us sceptical of a truth norm that features the normative notion of correctness. I don’t think the notion of correctness has a sufficient pre-theoretical intuitive basis to decide whether or not correctness comes by degree. ‘Correct’ seems too much of a term of an art for there to be a satisfactory answer to this question. Therefore, I don’t think we should give the aim of avoiding error more weight by claiming, like Wedgwood does, that suspending judgement has an intermediary degree of correctness.

Therefore, I’m going to opt for a different way of giving the aim of avoiding error greater weight, one that only appeals to the deontic vocabulary of ‘ought’ which I’ve predominantly been using. Deontic vocabulary like ‘ought’ definitely does not admit of degree. As we saw above, one either conforms to a deontic norm completely or one fails to do so completely – one cannot conform to a deontic norm to a certain degree. However, as I’ll outline in the next section, there is an alternative way to give the aim of avoiding error greater weight within a deontic framework.

3.2. My solution: a contrary-to-duty obligation to err on the side of caution

I’ll now outline how a truth norm account which holds that belief is subject to a deontic truth norm (like \( \textit{ought} \)) can give the aim of avoiding error greater weight. This, I’ll show, enables us to explain why one should follow the higher threshold evidential norm of \( \textit{SUFFICIENT EVIDENCE} \), and thus enables us to deal with the suspension of judgement problem.
I want to do this by claiming that the fundamental epistemic norms include not only a truth norm that gives weight to both aims, but also a norm that states that one of the two ways of failing to conform to this truth norm is preferable. In particular, this additional norm should state that with regards to any proposition \( p \), failing to conform to the truth norm with regards to \( p \) by not having a belief about whether \( p \) is preferable to failing to conform to the truth norm by having a false belief about whether \( p \). This would mean that the fundamental epistemic norms, as a whole, give greater weight to the aim of avoiding error.

The specific way I suggest we do this is by thinking that the fundamental epistemic norms include not only a truth norm, but also a contrary-to-duty obligation, which specifies what one ought to do if one fails to conform to that truth norm. Contrary-to-duty obligations are familiar from the practical case, though they have been predominantly discussed because of the puzzles they create in deontic logic. For an example, say one is under the following obligation:

**DO NOT KILL:** One ought not to kill anyone.

If one is under this obligation, there is also the question, however, of what one ought to do if one ends up violating it. As Chisholm puts it “[o]ur misdeeds generally create new duties” (1963, 33). In the case of **DO NOT KILL**, it seems that one is also under the following conditional contrary-to-duty obligation:

**KILL THEM GENTLY:** If one does kill someone, one ought to kill them gently.

I propose that an obligation of this kind can give the aim of avoiding error greater weight than the aim of believing the truth, and thereby solve the suspension-of-judgement problem. Specifically, while we should still maintain that belief is subject to a truth norm:

**OUGHT:** One ought to believe that \( p \) if and only if \( p \) is true.

We should also claim that belief is subject to the following conditional contrary-to-duty obligation:

---

4 I am grateful to Bernard Salow for suggesting this strategy to me.

5 The starting point for the deontic logic literature on contrary-to-duty obligations is Chisholm 1963.
ERR ON THE SIDE OF CAUTION: If one fails to conform to OUGHT with regards to p, one ought not to do so by believing falsely.

This contrary-to-duty obligation has the consequence that belief is unlike love, at least as described by Tennyson: “Tis better to have loved and lost than to never to have loved at all.” My claim is that belief is the opposite: it’s better to have never believed at all than to have believed falsely.

We should think of both OUGHT and ERR ON THE SIDE OF CAUTION as the two fundamental norms on belief which explain justification. OUGHT gives weight to both of the dual aims of believing the truth and avoiding error, and ERR ON THE SIDE OF CAUTION ensures that greater weight is given to the aim of avoiding error. And this means that the combination of these two norms can explain an evidential norm like SUFFICIENT EVIDENCE. In other words, if one is trying to conform to the combination of these two norms, one should follow SUFFICIENT EVIDENCE.

We can see this, I suggest, if we outline how the combination of OUGHT and ERR ON THE SIDE OF CAUTION improves on, a) OUGHT alone – which gives the dual truth goals equal weight – and b) permissive truth norms – which only give weight to the aim of avoiding error.

In contrast to permissive truth norms, the combination of OUGHT and ERR ON THE SIDE OF CAUTION does not give weight only to the aim of avoiding error. In this respect they improve on a permissive truth norm. A permissive truth norm only gives weight to the aim of avoiding error, and thus has the problematic result that one should suspend judgement about everything.

The combination of OUGHT and ERR ON THE SIDE OF CAUTION do not have the same problem. Suspending judgement about everything is not the best way to conform to these norms. This is because of ERR ON THE SIDE OF CAUTION’s status as a contrary-to-duty obligation. Because ERR ON THE SIDE OF CAUTION is a contrary-to-duty obligation, OUGHT takes precedence over it.

We can illustrate with the two norms DO NOT KILL and KILL THEM GENTLY. With these norms, the first norm takes precedence over the second. Therefore, one is not off the hook with regards to the combination of the two norms if one kills many people but

---

8 Canto 27, In Memorium A.H.H.
kills them all gently. Therefore, killing many people but killing them gently is not a good means of conforming with the combination of two norms. Likewise, one’s not off the hook with regards to the combination of OUGHT and ERR ON THE SIDE OF CAUTION if one suspends judgement about everything. Therefore, suspending judgement about everything is not a good means of conforming to the combination of OUGHT and ERR ON THE SIDE OF CAUTION.

The fact that trying to conform to the combination of OUGHT and ERR ON THE SIDE OF CAUTION does not, as a permissive truth norm does, involve suspending judgement about everything shows that the two norms do not only give weight to the aim of avoiding error. But the combination does give more weight to the aim of avoiding error, and that’s the reason why it improves on OUGHT alone.

The problem with OUGHT alone was that it could only justify giving the dual aims of believing the truth and avoiding error equal weight. This meant that it could only explain an evidential norm like MORE THAN 50-50, and not one like SUFFICIENT EVIDENCE. This is because suspending judgement about whether p in cases in which p is only slightly more likely than its negation given one’s evidence – as following SUFFICIENT EVIDENCE – looks like it privileges the aim of avoiding error over the aim of believing the truth.

Adding ERR ON THE SIDE OF CAUTION gives the aim of avoiding error more weight, because it states that one of the two ways of failing to conform to OUGHT – failing to have a true belief about whether p – is preferable to the other – having a false belief about whether p. This deals with the problem OUGHT had in explaining SUFFICIENT EVIDENCE. If the aim of avoiding error has greater weight, then one should not believe that p if one’s evidence only makes p slightly more likely than its negation, i.e. one should not follow MORE THAN 50-50. Rather one should believe that p only if one’s evidence supports p above some higher threshold of evidential support. Taking this latter option looks like following SUFFICIENT EVIDENCE. Therefore, whereas OUGHT alone failed to explain SUFFICIENT EVIDENCE, it looks like the combination of OUGHT and ERR ON THE SIDE OF CAUTION can succeed in explaining it.

And if SUFFICIENT EVIDENCE can be explained by the combination of OUGHT and ERR ON THE SIDE OF CAUTION, then that provides us with an answer to the suspension-of-judgement problem. According to my account, it’s still the case that it’s rationally impossible to conform to OUGHT by suspending judgement about whether p.
But the means one should take to try and conform to the combination of OUGHT and ERR ON THE SIDE OF CAUTION involve following an evidential norm — SUFFICIENT EVIDENCE — that involves suspending judgement in many cases, i.e. cases in which one doesn’t have sufficient evidence either way. This essentially is Engel’s response to the suspension-of-judgement problem, but with a pair of fundamental epistemic norms that actually can succeed in explaining the requisite evidential norm.

I’ll now deal with a possible objection to this account. One might worry that the addition of ERR ON THE SIDE OF CAUTION is ad hoc, i.e. one might argue that ERR ON THE SIDE OF CAUTION is not well motivated because I haven’t given any independent reason to think that belief is subject to this kind of norm; I’ve only argued that it would, if added to OUGHT, allow us to successfully explain an evidential norm like SUFFICIENT EVIDENCE.

I accept that I haven’t given any independent motivation to think that belief is subject to ERR ON THE SIDE OF CAUTION. But given the methodology I outlined and defended in Chapter 2, this is unproblematic. As I claimed there, I do not think we should start with a truth norm which we think belief is subject to for independent reasons, and then go on to use that truth norm to explain justification. Rather, we should think that belief is subject to a truth norm because belief being subject to such a norm can do justice to our pre-theoretical intuitions about justification.

From this point of view, what we’ve shown in this chapter is that belief cannot just be subject to a truth norm like OUGHT, because such a norm on its own fails to explain justification — in particular it fails to explain how one can be unjustified in believing that p when one’s evidence supports p more than not-p, but only slightly. According to our methodology, if the addition of ERR ON THE SIDE OF CAUTION results in an account that succeeds in explaining justification, by explaining why one counts as unjustified in such cases, that by itself is reason to think that belief is subject to this additional fundamental epistemic norm. Therefore, although ERR ON THE SIDE OF CAUTION is designed specifically to deal with the issue at hand, it is not ad hoc in any sense that’s damaging to the resulting account.

4. Conclusion

In this chapter we’ve seen that suspension of judgement creates a more fundamental problem for the truth norm account of justification than has hitherto been realised. This
problem is not answered simply by appealing to the fact that conforming a truth norm requires following a norm that prescribes believing that p if and only if one has enough evidence for p (SUFFICIENT EVIDENCE). This is because this response ignores the possibility of lower-threshold evidence norms (such as MORE THAN 50-50) which prescribe suspending in far fewer cases. In order to justify a higher-threshold evidence norm, we need to give the aim of avoiding error greater weight. And we should do that, I’ve suggested, by augmenting the basic truth norm picture with a contrary-to-duty obligation to err on the side of caution.
As I outlined in Chapter 1, the truth norm account of justification is a particular version of the more general claim, which I called ‘veritism’, that true belief is the fundamental goal of epistemic justification. This chapter discusses a problem with veritism in general. Particular examples are alleged to create problems for veritism: examples in which having a belief that looks prima facie unjustified is a means to the end of having a large number of other true beliefs. I’ll call examples of this kind ‘Firth cases’ since they were first discussed by Roderick Firth (1981). Many philosophers have claimed that Firth cases pose problems for some forms of veritism (Conee 1992, 249–51; Fumerton 2001, 54–56; Jenkins 2007, 37–38; Littlejohn 2012, 80–82; Greaves 2013, especially 949–51), but in a series of recent articles Selim Berker has argued that Firth cases should lead us to reject veritism tout court (2013a; 2013b; 2015).

This conclusion, I’ll argue, is unwarranted. The truth norm account of justification is a form of veritism which does not have a problem dealing with Firth cases. This is because, as I’ll outline, the truth norm account is a version of what I’ll call ‘proposition-relative veritism’. Proposition-relative veritism claims that a belief that \( p \) is justified if and only if having that belief is a good means to having a true belief that \( p \). According to this version of veritism, the fact that believing that \( p \) is a means to the end of having other true beliefs, as happens in Firth cases, makes no difference to whether or not one would be justified in believing that \( p \).

I’ll proceed as follows. In §1, I’ll describe the problem Firth cases raise for veritism. In §2, I’ll distinguish between ‘proposition-neutral’ and ‘proposition-relative’ versions of veritism and argue that Firth cases are unproblematic for the latter. In §3, I’ll discuss and respond to two objections Berker raises for proposition-relative veritism. In responding to one of these objections, I’ll claim that we need amend the basic truth norm picture with an additional causal independence condition on justification.

1. Veritism and Firth cases

The definition of veritism I started with in Chapter 1, which I took from Laurence BonJour (1985, 7–8), was the claim that having justified beliefs is a means to the end of
having true beliefs:

VERITISM: A belief is justified if and only if having that belief is a good means to the end of having true beliefs.

As I said in Chapter 1, the truth norm account of justification is a specific version of veritism. However, it will serve our purposes in this chapter to discuss veritism more generally rather than the truth norm account of justification in particular. Firstly, this is because Berker thinks Firth cases are a problem for any form of veritism. Secondly, starting with veritism in its most general form enables us to see most clearly how this claim of Berker’s is incorrect, i.e. it is through distinguishing different modifications of this basic veritist formula that we can understand which versions of veritism genuinely are vulnerable to Firth cases and which versions are not.

With veritism defined, we can now outline the problem raised by Firth cases. We can illustrate this, as Berker does, with the following Firth case given by Richard Fumerton (2001, 55). Say a scientist is seeking to get a grant from a religious organisation that only gives grants to true believers. This scientist does not believe that God exists; he believes there is no evidence that God exists, and that anyone who believes God exists is irrational. Knowing himself to be a terrible actor, this scientist knows that he'll only get the grant if he genuinely believes that God exists. Therefore, if this scientist believes that God exists, that will greatly increase his chances of getting the grant, which will allow him to continue and expand his research. And continuing and expanding his research will lead him to have a large number of new true beliefs.

This case is problematic, Berker claims, because veritism has the counterintuitive consequence that the scientist would be justified in believing that God exists. This is because believing that God exists would be, for the scientist, a good means to the end of having true beliefs. In fact, believing that God exists would be, Berker claims, a better means to the end of having true beliefs than not believing that God exists (2013a, 364). This is because although believing that God exists is not, ex hypothesi, a good means to the end of him having a true belief that God exists, it is a good means to the end of him having many other true beliefs, i.e. true beliefs about the subject matter of his scientific research. With regards to the goal of having true beliefs, this case involves the prospect of a trade-off or, as Berker puts it, a 'sacrifice for the greater epistemic good' (2013a, 364), i.e. not achieving that goal with respect to one proposition, but achieving that goal with regards
to a large number of other propositions. And Berker’s claim is that a believer who makes these kinds of trade-offs is doing better with regards the goal of having true beliefs than a believer who does not.

If Berker is right about this, then that casts doubt on veritism. Intuitively, the scientist in the above example would not be epistemically justified in believing that God exists. Therefore, the veritist account of justification looks like it gets the wrong result in this example.

This is the basic structure of Berker’s argument against veritism. All forms of veritism, Berker claims, will have the wrong results in some cases like the example of the scientist. In defending this general claim, Berker considers and rejects a number of modified forms of veritism that do not have the consequence that the scientist would be justified in believing that God exists. These include restricting the truth goal to the goal of having true beliefs now (Berker 2013a, 367), and claiming that a justified belief is one that tends to result in the furthering of the truth goal, rather than one that just does so on a particular occasion (Berker 2013b, 369–70). Berker rejects these responses because he claims that we can construct further examples involving the same kind of trade-offs that show that even such modified versions of veritism are untenable (Berker 2013a, 369–77; 2013b, 369–78).

I’m not going to dispute Berker’s rejections of possible responses (with the exception of his objections to a proposition-relative form of veritism, which I’ll discuss in §3). This is because I want to discuss what Berker takes to be the fundamental problem with veritism.1 Berker claims all forms of veritism ignore what he calls ‘the epistemic separateness of propositions’:

The more general point is this: when determining the epistemic status of a belief in a given proposition, it is epistemically irrelevant whether or not that belief conduces (either directly or indirectly) toward the promotion of true belief and the avoidance of false belief in other propositions beyond the one in question. (Berker 2013a, 365)

Berker intends this problem with veritism to be analogous to the claim, most famously made by Rawls, that consequentialism in practical ethics “does not take seriously the

---

1 Most of the responses to Berker to date have tried to defend particular forms of veritism. See e.g. Goldman’s (2015) and Ahlstrom-Vij and Dunn’s (2014) defences of reliabilism against Berker’s criticisms.
distinction between persons” (Rawls 1999, 26). This is often supposed to be manifest by the fact that certain forms of consequentialism prescribe making trade-offs that look prima facie morally wrong because they involve violating someone's rights. Such cases include the case of a transplant surgeon who can save five people’s lives by killing one healthy person and taking his organs (Thomson 1976, 206), or the case in which the authorities can prevent riots and lynchings by punishing an innocent person for a crime (McCloskey 1957, 468–69). According to Berker, veritism inevitably prescribes making epistemic versions of these kinds of trade-offs because it ignores the separateness of propositions (2013a, 365).

I’m going to dispute this claim of Berker's. Veritism does not inevitably ignore the epistemic separateness of propositions, and it does not inevitably have the counterintuitive consequence that believers who make trade-offs for the epistemic greater good would count as justified. In fact, as I’ll show in the next section, a plausible version of veritism – of which my truth norm account is an instance – both entails that believers who make such trade-offs would not count as justified, and explains the epistemic separateness of propositions.

2. An answer: proposition-relative veritism

We can see how veritism can deal with Firth cases if we distinguish between a proposition-neutral and a proposition-relative understanding of the view. Proposition-neutral veritism is the claim that having a justified belief is a good means to the end of having true beliefs in general. Proposition-relative veritism is the claim that having a justified belief that p is a good means to the end of having a true belief that p, i.e. having a true belief about that very proposition. My claim is that Firth cases create problems for a proposition-neutral conception of veritism, but not for a proposition-relative conception.

Let’s define proposition-neutral veritism as follows:

**PROPOSITION-NEUTRAL VERITISM** A belief is justified if and only if having that belief is a good means to the end of having true beliefs in general.

The following passage from William Alston suggests this conception of veritism:

> Epistemic evaluation is undertaken from what we might call the “epistemic point of view.” That point of view is defined by the aim at maximizing truth and minimizing falsity in a large body of beliefs. … [O]ur central cognitive aim is to amass a large body of beliefs with a
favorable truth-falsity ratio. For a belief to be epistemically justified is for it, somehow, to be awarded high marks relative to that aim. (Alston 1985, 59)

This suggests a propositional-neutral version of the truth goal because it states that a particular belief is justified if it has ‘high marks’ relative to the aim of having a favourable truth-falsity ratio in one’s beliefs in general.

On the other hand, we can define proposition-relative veritism as follows:

**PROPOSITION-RELATIVE VERITISM**: A belief that p is justified if and only if having that belief is a good means to the end of believing that p if and only if p is true.

The following passage from Wedgwood explicit puts forward this conception of veritism and distinguishes it from the proposition-neutral version:

[T]here is a separate goal, for each particular proposition p that one actually considers – roughly, the goal of believing this proposition p if and only if p is true – not the general goal of amassing as many true beliefs as possible. (Wedgwood 2008, 5, fn.8)

According to this conception of veritism, the only aim or goal that is relevant to whether or not a belief that p is justified is the goal of believing that p if and only if p is true. Whether or not a belief that p is justified depends on whether having that belief is a good means to that end, according to proposition-relative veritism, and not in virtue of how it contributes to any more general aim of having beliefs with a favourable truth-falsity ratio, as on the proposition-neutral conception.

It should be clear that the truth norm account of justification is a proposition-relative version of veritism. This is because all of the truth norms we’ve considered are ones which claim that whether or not one ought to believe the proposition that p depends on whether or not the proposition that p is true. Because of this, it’s no surprise that we found a clear statement of proposition-neutral veritism in Wedgwood, who’s a proponent of the truth norm account.

This distinction matters because only proposition-neutral veritism has a problem dealing with Firth cases. Proposition-relative veritism does not have the same problems with these cases.

We can illustrate this with the scientist example. For the scientist, is believing that God exists a good means to achieving the goal of having true beliefs? It surely is on the proposition-neutral way of understanding that goal. Believing that God exists is a good
way for the scientist to achieve the goal of, to use Alston’s words, ‘amassing a large body of beliefs with a favorable truth-falsity ratio’. And so, according proposition-neutral veritism, then this scientist will count as justified. But, as we saw earlier, this looks wrong – this scientist would not be justified in believing that God exists – and so we should reject proposition-neutral veritism.

But proposition-relative veritism does not have the same problem with this case. Believing that God exists is not a good way for the scientist to believe that God exists if and only if it’s true that God exists. On the proposition-relative conception of veritism, the fact that believing that God exists may lead to the scientist having other true beliefs is irrelevant to the question of whether the belief that God exists is justified.

In this way, proposition-relative veritism does not suffer from what Berker claims is the fundamental defect in veritism. It does not ignore the epistemic separateness of propositions, i.e. the fact that “when determining the epistemic status of a belief in a given proposition, it is epistemically irrelevant whether or not that belief conduces … toward the promotion of true belief and the avoidance of false belief in other propositions beyond the one in question” (Berker 2013a, 365).

In fact, proposition-relative veritism can explain this irrelevance. That believing that p may lead to true beliefs in other propositions is epistemically irrelevant because this fact does not necessarily count in favour of the proposition that p being true, as in Firth cases.

Therefore, Berker is wrong to think that all forms of veritism must ignore the epistemic separateness of propositions, and thereby be committed to the counterintuitive consequence that believers who make ‘sacrifices for the greater epistemic good’ would be justified. Proposition-relative veritism – of which the truth norm account of justification is a version – does not have this consequence. Therefore, although Firth cases should lead us to reject proposition-neutral veritism, they should not lead us to reject veritism tout court.

3. Objections and replies

I'll now discuss two objections Berker makes against proposition-relative veritism. Firstly, Berker argues that making truth goals proposition-relative does not deal with Firth cases because one still has to weigh distinct proposition-relative truth goals against each other. I'll respond to Berker, in §3.1, by arguing that a sensible proposition-relative
veritism involves *denying* that one needs to weigh distinct proposition-relative truth goals against each other.

Secondly, Berker argues that proposition-relative veritism is still vulnerable to modified versions of Firth cases (or something like them). In particular, he claims that proposition-relative veritism has the wrong results in cases of self-fulfilling beliefs, i.e. beliefs that cause their own truth. I'll argue, in §3.2, that self-fulfilling beliefs don't give us reason to reject veritism _tout court_ (as Berker does). I'll argue they can be dealt with in a different way, amending proposition-relative veritism with an additional _causal independence_ condition.

### 3.1. The problem of weighing proposition-relative goals against each other

Berker argues that the proposition-relative veritist way of dealing with Firth cases is undermined because, he claims, one still has to weigh proposition-relative goals against each other. This is because of the fact that “one can come to believe that p at the same time as one comes to, or continues to, believe that q (where <p> and <q> are distinct propositions)” (Berker 2013a, 376–77). According to proposition-relative veritism, I have the _both_ the aim of believing that p if and only if p is true _and_ the distinct aim of believing that q if and only if q is true. And this means, Berker claims, that in the case in which I come to believe that p at the same time as I believe that q, I have to _weigh_ my aim of believing that p if and only if p is true _against_ my aim of believing that q if and only if q is true. As Berker puts it, “even proposition-relative veritism must weigh the furtherance or frustration of our epistemic aims with regard to one proposition against the furtherance or frustration of our epistemic aims with regard to other propositions” (Berker 2013a, 377).

Berker’s objection is that since one has to weigh distinct proposition-relative truth goals against each other, the problem raised by Firth cases simply reappears. In other words, Berker is claiming that, in a Firth case, the fact that one won’t achieve one proposition-relative truth goal will inevitably be outweighed by the fact that one will achieve many other proposition-relative truth goals. To illustrate, proposition-relative veritism is going to have the result, according to Berker, that the scientist would be justified in believing that God exists, because in such a case the fact that one will _not_ achieve the aim of believing that God exists just in case God exists will inevitably be outweighed by achieving the multiple aims of believing that p if and only if p is true, of
believing that q if and only if q is true, of believing that r if and only if r is true, … (where p, q, r, … are propositions about some matters the scientist intends to research).

I don’t think Berker has raised a genuine problem with proposition-relative veritism here, because I think he’s wrong to think that proposition-relative veritism requires believers to do the kind of ‘weighing’ he envisages.

The proposition-relative veritist position is that what determines whether a belief that p is justified is whether believing that p is a good means to the end of believing that p if and only if p is true; whether this results in one failing to meet the separate epistemic aim of believing that q if and only if q is true is irrelevant to whether or not a belief that p is justified. This is, in effect, to deny that in considering whether my belief that p is justified I weigh my aim of believing that p if and only if p is true against my other aims to believe that q if and only if q is true, to believe that r if and only if r is true, and so on; all that matters is whether believing that p is a good means to the end of believing that p if and only if p is true.

Therefore, Berker is wrong to claim that ‘even proposition-relative veritism must weigh the furtherance or frustration of our epistemic aims with regard to one proposition against the furtherance or frustration of our epistemic aims with regard to other propositions’. According to proposition-relative veritism, one does not, contra Berker, have to weigh one’s aim of believing that p if and only if p is true against a separate aim of believing that q if and only if q is true, in the sense of considering whether to sacrifice one aim to meet the other. The effects of believing that p on one’s other epistemic aims is irrelevant to whether one is justified in believing that p.

Why does Berker claim that the proposition-relative veritist must weigh distinct proposition-relative truth goals against each other, even though the account actually has the result, as I’ve shown, that this kind of weighing is irrelevant? I can think of two possible rationales. But I’ll suggest that neither of these rationales in fact supports such a claim.

Firstly, Berker may think that the proposition-relative veritist must weigh distinct proposition-relative truth goals against each other because of the analogy between veritism and consequentialism. Berker often is an epistemic form of consequentialism, because it holds that whether one is justified is determined by whether having one’s beliefs is a good or bad way of producing certain outcomes, i.e. having true beliefs (2013a, 341–50; 2013b, 365–69). And I suspect that Berker, in claiming that the proposition-
relative veritist must weigh distinct proposition-relative truth goals against each other, may be assuming that veritism must share a particular feature of practical consequentialism.

The feature of practical consequentialism I have in mind is its emphasis on aggregation. Practical consequentialism holds that what one ought to do depends not on a particular person’s interests but on the aggregated set of everyone’s interests. To illustrate, say I have the job of organising the distribution of donated organs, and consequentialism prescribes giving the one liver I have to Jane, who’s young, rather than to John, who’s very old, because Jane will get more benefit from it. Consequentialism does not prescribe giving the liver because of how that will affect Jane’s interests taken in isolation. It prescribes giving the liver to Jane because of how much weight Jane’s interests have in the aggregated set of everyone’s interests – i.e. because the amount of happiness (or preference satisfaction or quality-adjusted life years) Jane will receive from the liver results in a higher aggregated total than what would result from giving the liver to John. In this way, according to consequentialism, Jane’s interests explain why I ought to give her the liver only because of how much weight they have in the utilitarian calculus as a whole.

Berker may think that proposition-relative veritism, in virtue of being a form of epistemic consequentialism, must share this feature of practical consequentialism, i.e. that proposition-relative veritism must claim that what one ought to believe is explained by how the many distinct proposition-relative truth goals contribute to an aggregated set of truth goals. And perhaps Berker is assuming that what one ought to believe, on the proposition-relative veritist picture, is whatever will further the what results in the largest aggregated set of achieved truth goals, i.e. the largest number of true beliefs. If this were the case, then Berker would be correct that even the proposition-relative veritist must weigh distinct proposition-relative truth goals against each other, because that would be involved in figuring out how one can achieve as many proposition-relative truth goals as possible.

However, if these are Berker’s reasons for thinking that the proposition-relative veritist must weigh distinct proposition-relative truth goals against each other, then they are not good reasons. This is because a sensible proposition-relative veritist should, I suggest, simply deny that proposition-relative truth goals aggregate. A sensible proposition-relative veritism should claim that there are only a number of separate goals
for each proposition – i.e. the goal to believe that $p$ if and only if $p$ is true, the goal to believe that $q$ if and only if $q$ is true, etc. – and claim that what one ought to believe is is not determined by an aggregated set of these goals. In fact, denying that proposition-relative truth goals aggregate really is what distinguishes proposition-relative veritism from proposition-neutral veritism in the first place. Therefore, Berker would be wrong if he thinks proposition-relative veritism – in virtue of being a form of epistemic consequentialism – must claim that proposition-relative truth goals aggregate. This is, therefore, one place in which Berker’s analogy between veritism and consequentialism breaks down.

A second possible rationale is that Berker thinks that the proposition-relative veritist must weigh distinct proposition-relative truth goals against each other in order to explain what goes on when one weighs different hypotheses against each other.

We can find an example of this kind of weighing in Hume’s discussion of miracles. Say a friend of mine, Mary, who I’ve known in the past to be reliable and honest, tells me that a man has risen from the dead. Working out what I would justified in believing in such a case would involve weighing the hypothesis that a man has risen from the dead against other competing hypotheses, including the hypothesis that Mary is either lying or deceived. If I follow Hume, I’m liable to think that, given Mary is telling me that something has happened “that has never been observed in any age or country” (Hume 1748, 90), it’s more likely to be true that Mary is either lying or deceived, and conclude that I would not be justified in believing that a man has risen from the dead.

Berker may think that for proposition-relative veritism to explain this kind of weighing – the weighing of competing hypotheses against each other – it needs to appeal to the kind of weighing he envisages – the weighing of distinct proposition-relative truth goals against each other. What I have in mind is the idea that for proposition-relative veritism to explain why the competing hypothesis that Mary is lying or deceived is relevant to whether I would be justified in believing that a man has risen from the dead, we need to appeal to weighing truth goals against each other, i.e. that what goes on when I weigh the hypothesis that Mary is either lying or deceived against the hypothesis that a man has risen from the dead is that I weigh my proposition-relative aim of believing that Mary is either lying or deceived if and only if that’s true against my distinct aim of believing that a man has risen from the dead if and only if that’s true.

If this is Berker’s reason for thinking that the proposition-relative veritist must
weigh distinct proposition-relative truth goals against each other, then it is also not a good reason. This is because proposition-relative veritism can easily explain why the hypothesis that Mary is either lying or deceived is relevant to whether I'm justified in believing that a man has risen from the dead, and without appealing to the weighing of distinct proposition-relative truth goals against each other.

This is because proposition-relative veritism claims that, in considering whether one would be justified in believing that \( p \), the only factors one needs to consider are factors that bear on the truth or falsity of \( p \) (because only those factors bear on whether or not having that belief is a good means to believing that \( p \) if and only if \( p \) is true).

This, by itself, suffices to explain why the hypothesis that Mary is either lying or deceived is relevant to whether or not I'm justified in believing that a man has risen from the dead. This is because one factor that bears on whether or not it's true that a man has risen from the dead is whether or not it's true that Mary is lying or deceived. This is because these two propositions are inconsistent. If one is true, the other is false. And if I'm considering whether or not believing that a man has risen from the dead is a good means of believing that a man has risen from the dead if and only if it's true, I'd better consider whether any salient inconsistent hypotheses are likely to be true.

This is why weighing competing hypotheses is relevant to justification on the proposition-relative veritist picture. And it is an explanation that does not need to appeal to the weighing of distinct proposition-relative truth goals against each other. In the above example, I do not weigh my aim of believing that a man has risen from the dead just in case that’s true against my aim of believing that Mary is either lying or deceived just in that case that’s true.

In summary, if Berker were correct in claiming that the proposition-relative veritist must weigh distinct proposition-relative truth goals against each other, then moving to a proposition-relative conception of the truth goal would not deal with Firth cases. However, proposition-relative veritism actually involves denying that one has to do this kind of weighing. And neither of the two rationales we’ve covered gives any good reason to think that the proposition-relative veritism must involve this kind of weighing.

3.2. The problem of self-fulfilling beliefs?

Berker also argues that proposition-relative veritism is still vulnerable to Firth cases (or something like them). He argues that there are cases in which believing that \( p \) is a good
means of believing that p if and only if p is true – i.e. a good means of furthering the proposition-relative truth goal – but in which believing that p looks unjustified. He claims that self-fulfilling beliefs – beliefs that cause their own truth – are examples of such cases (2013b, 376–77).

Berker illustrates this with the example of Jane, who has been diagnosed with an illness that is likely to be terminal unless she believes that she will recover from it. If Jane does not believe that she will recover, she will only have a 10% chance of recovering. But if Jane does believe that she will recover, that will raise her chances of recovering to 90% (Berker 2013b, 376).

How does this example make trouble for proposition-relative veritism? It does so because believing that she will recover is, for Jane, a good means of believing that she will recover just in case she’ll recover – i.e. it is a good means to the end of having a true belief about that very proposition. Therefore, believing that she’ll recover is, for Jane, a good way to further the proposition-relative truth goal. But Berker insists that “Jane’s belief is not for that reason epistemically justified” (2013b, 376). For this reason Berker concludes that the shift to a proposition-relative conception of the truth goal does not deal with the fundamental problem with veritism.

One option the veritist could take here in response to Berker would be to accept the consequence that self-fulfilling beliefs are justified (cf. James 1897, chap. 1; Velleman 1989, chap. 2). I’m not going to consider this option since I agree with Berker that self-fulfilling beliefs are not epistemically justified (though, for a response to those who claim that they are, see Langton 2004). I also agree that the shift to a proposition-relative conception of veritism, although it deals with all other Firth cases, can’t by itself explain why one would be unjustified in forming a belief just because it’s self-fulfilling. But I disagree that this should lead us to reject veritism tout court. This is because proposition-relative veritism can deal with self-fulfilling beliefs in a different way.

I am going to suggest that a proposition-relative veritist should, given the broad motivations for veritism, claim that there is a further necessary condition on justified belief. Specifically, a veritist should introduce a causal independence on justification.

This is not an ad hoc addition to veritism. A good veritist should, I suggest, be committed to such a causal independence condition. This is because the same motivation for veritism also supports a causal independence condition. The broad motivation for veritism is that beliefs are the kind of things that are supposed to represent how the
world is independently of how the believer takes it to be. It is suggested by defenders of veritism that this feature of belief warrants the veritist claim that truth is the fundamental goal of justification (see, e.g. Velleman 2000; Wedgwood 2002).

This feature of belief also, I suggest, warrants thinking that there is an additional causal independence condition on justification. In fact, self-fulfilling beliefs show that the proposition-relative veritist truth goal cannot by itself explain this feature. This is because having a self-fulfilling belief is a good means of furthering the truth goal, but is not a good means of representing how things are independently of how the believer takes them to be. Adding a causal independence condition gives us an account of justification that fully accounts for the fact that beliefs are supposed to represent how the world is independently of the believer takes it to be.

Therefore, I suggest we add the following causal independence condition to proposition-relative veritism. How might we formulate such a condition? A first pass would be as follows:

**CAUSAL INDEPENDENCE**: One has a justified belief that \( p \) only if the truth of \( p \) is not causally dependent on the fact that one believes that \( p \).

I should note that a fact \( A \) being ‘causally dependent’ on another fact \( B \), as I’m using the term, means more than just \( B \) having some causal influence on \( A \); in my terms, if \( A \) is causally dependent on \( B \), then if \( B \) were to not to be the case, then \( A \) would not be the case. This means that **CAUSAL INDEPENDENCE** does not rule out as justified cases in which you truly believe that \( p \) on the basis of good evidence, but your belief also (unbeknownst to you) increases the likelihood of its own truth to some degree. Such a belief would not be ruled out as justified by causal independence because its truth is not solely dependent on the fact that you have it.

I should also note that **CAUSAL INDEPENDENCE** does not claim that in order for me to have a justified belief that \( p \), the truth of \( p \) can’t be causally dependent on any of my beliefs; only that the truth of \( p \) cannot be dependent on my belief that \( p \).

So how does **CAUSAL INDEPENDENCE** rule out self-fulfilling beliefs from being justified? Self-fulfilling beliefs do, as I admitted earlier, meet our original proposition-relative veritist condition on justification. For example, believing that she’ll recover is a good means, for Jane, of believing that she’ll recover if and only if it’s true that she’ll recover. But such a self-fulfilling belief does not meet the causal independence condition.
The truth of whether Jane will recover is causally dependent on the fact that Jane believes that she’ll recover. Therefore, a self-fulfilling belief like Jane’s does not count as justified if we add this causal independence condition. And the same will be the case with all cases of evidentially unsupported self-fulfilling beliefs. Since, *ex hypothesi*, the truth of such beliefs is causally dependent on the fact the believer holds them, no such belief will count as justified according to this modified account.

Therefore, we should amend proposition-relative veritism with the causal independence condition. Although self-fulfilling beliefs do show that Berker was correct to claim that Firth cases remain for proposition-relative veritism, it can be modified so as to deal with them. And once we’ve done this, it doesn’t seem that there will be any Firth cases left, i.e. cases in which believing that p is both a good means believing that p if and only if p is true *and* in which p’s truth is causally independent of one’s belief that p, but in which one looks prima facie unjustified in believing that p.

4. Conclusion

Veritism does not, in the end, face an issue with Firth cases. I’ve outlined a form of veritism which does not have a problem in dealing with these sacrifices for the greater epistemic good. However, I’ve admitted that a defensible version of veritism also has to involve a causal independence condition.

Given the truth norm account is a version of veritism, this chapter has also, like the last two, shown another amendment is necessary in order to give a defensible explanation of justification in terms of a truth norm. A truth norm will, as I said earlier, always result in a proposition-relative version of veritism, so it will deal with most Firth cases. But to deal with the full range of Firth cases, we also need to add a causal independence condition to the truth norm account of justification.
CONCLUSION

Part One of this thesis laid the groundwork for an epistemological argument for the claim that belief is subject a truth norm. If a truth norm could explain how our pre-theoretical intuitions about justification all hang together, that would be sufficient reason to conclude that belief is in fact subject to a truth norm.

This epistemological argument for a truth norm had the promise of being much more convincing than any psychological argument for a truth norm, given the numerous flaws such psychological arguments suffer from. However, this epistemological argument is only going to work if a truth norm can genuinely play the role of explaining justification.

Part Two of this thesis followed through on this promise. Despite there being a number of problems that have been raised for an explanation of justification in terms of a truth norm, I've defended a version of the truth norm account of justification that can deal with this problems, and thereby succeed in explaining justification. In order to do this, I've made three main modifications to the basic truth norm picture.

Firstly, Chapter 5 showed that the believing-all-the-truths problem should lead us accept a conditional truth norm; in particular, a truth norm that is only in effect if one has some doxastic attitude towards the proposition in question:

DOXASTIC ATTITUDE: If one has some doxastic attitude towards p, one ought to believe that p if and only if p is true.

Only this formulation of a truth norm was able to escape the problematic implications of a prescription to believe all the truths, but not thereby become incapable of explaining justification.

Secondly, Chapter 6 showed that the suspension-of-judgement problem should lead us to accept that there is a contrary-to-duty obligation to err on the side of caution. In Chapter 6, I outlined this claim in relation to OUGHT. Therefore, from the point of view of the thesis as a whole, we should adopt a contrary-to-duty obligation that is relative to DOXASTIC ATTITUDE, since that is the truth norm I am defending:

ERR ON THE SIDE OF CAUTION-DA: If one fails to conform to DOXASTIC ATTITUDE with regards to p, one ought not to do so by believing falsely.
A contrary-to-duty norm like this gives the aim of avoiding error greater weight than the aim of believing the truth, and it is only by doing that that the truth norm account can explain an evidence norm that prescribes believing something if and only if one has sufficient evidence for it. And a truth norm needs to explain this kind of evidence norm in order to succeed in explaining justification.

Thirdly, Chapter 7 showed that in order to deal with Firth cases, we should add a causal independence condition on justification to the basis truth norm picture.

CAUSAL INDEPENDENCE: One has a justified belief that p only if the truth of p is not causally dependent on the fact that one believes that p. A truth norm by itself – in virtue of being a proposition-relative form of veritism – can deal with most Firth cases, i.e. it can explain why one would not in most cases count as justified by making a sacrifice for the greater epistemic good. However, self-fulfilling beliefs provided a recalcitrant class of problem cases. Amending the basis truth norm picture by adding the causal independence condition, however, rules out self-fulfilling beliefs from counting as justified.

Answering these three problems has, in this way, shown the contours of a defensible explanation of justification in terms of a truth norm. A truth norm with these three modifications am succeed in explaining how our pre-theoretical intuitions about justification all hang together.


Firth, R. 1981. “Epistemic Merit, Intrinsic and Instrumental.” *Proceedings and Addresses of*


Press.


