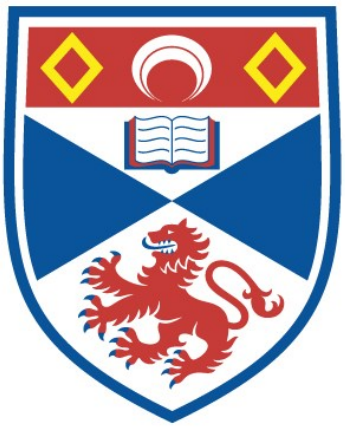
A defence of the truth norm of belief

Xintong Wei

A thesis submitted for the degree of PhD
at the
University of St Andrews



2022

Identifier to use to cite or link to this thesis:

DOI: https://doi.org/10.17630/sta/730

Candidate's declaration

I, Xintong Wei, do hereby certify that this thesis, submitted for the degree of PhD, which is approximately 73,000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for any degree. I confirm that any appendices included in my thesis contain only material permitted by the 'Assessment of Postgraduate Research Students' policy.

I was admitted as a research student at the University of St Andrews in February 2017.

I received funding from an organisation or institution and have acknowledged the funder(s) in the full text of my thesis.

16/8/2021

Date Signature of candidate

Supervisor's declaration

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree. I confirm that any appendices included in the thesis contain only material permitted by the 'Assessment of Postgraduate Research Students' policy.

16/8/2021

Date Signature of supervisor

Permission for publication

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand, unless exempt by an award of an embargo as requested below, that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that this thesis will be electronically accessible for personal or research use and that the library has the right to migrate this thesis into new electronic forms as required to ensure continued access to the thesis.

I, Xintong Wei, confirm that my thesis does not contain any third-party material that requires copyright clearance.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

Printed copy

No embargo on print copy.

Electronic copy

No embargo on electronic copy.

16/8/2021

Date Signature of candidate

16/8/2021

Date Signature of supervisor

Underpinning Research Data or Digital Outputs

Candidate's declaration

I, Xintong Wei, hereby certify that no requirements to deposit original research data or digital outputs apply to this thesis and that, where appropriate, secondary data used have been referenced in the full text of my thesis.

16/8/2021

Date Signature of candidate

Abstract

For belief, the standard of correctness is truth. Truth is said to be the norm of belief. This thesis aims to address some recent challenges relating to various aspects of the truth norm of belief on a reason-based normative framework. On a reason-based normative framework, normativity is a matter of (normative) reasons. Reasons are facts that count in favour of a response, grounded in value-based, desire-based, constitutive, or practice-based facts. Reasons are the sort of things that we can respond to and they have weights in making a claim on us when we stand in a relation to them. Three anti-normativist challenges arise in recent discussion. First, the grounding challenge concerns whether the reason to believe truly (correctly) and to refrain from believing falsely (incorrectly) can be properly grounded. Second, the guidance challenge concerns whether the reason to believe truly and to refrain from believing falsely is something we can respond to in our belief-formation. Third, the weighting challenge concerns whether the reason to believe truly and to refrain from believing falsely has any weight in making a claim on us regarding what we ought to/may believe. In this thesis, I offer novel responses to all three challenges. I develop and defend a practice-based, variantist account of the truth norm, according to which, the truth norm of belief is grounded in a justified social practice, guides our belief-formation on a reason-responsive model of epistemic guidance, and makes varying claims on us regarding what we ought to/may believe depending on the circumstances.

Acknowledgements

I would like to thank, first and foremost, Philip Ebert for his invaluable advice and steady support throughout my postgraduate studies. Philip always has a sympathetic ear to my ideas and has so generously given his time and energy to guide me through every stage of my research. I consider myself incredibly lucky to have been one of his doctoral students. He is not only an amazing mentor but also a great friend. I cannot imagine a better supervisor than him and this project would not have been possible without his support. I am also deeply grateful to my secondary supervisor, Crispin Wright, for his insightful feedback, support, and encouragement. His work on truth, self-knowledge and epistemology has greatly influenced my thinking.

I have benefited enormously from being a member of the research projects *Knowledge Beyond Natural Sciences* (2017-2019) and *Varieties of Risk* (2020-2021). I had the opportunity to work with Carrie Ichikawa Jenkins, Jonathan Jenkins Ichikawa, Sonia Roca-Royes, Peter Sullivan, Alan Miller, Alisa Mandrigin, Giacomo Melis, Giovanni Merlo, Martin Smith, Antony Duff, Sandra Marshall, Claire Field, Giada Fratantonio, Paul Colan, Sam Symons, Indrek Lobus, Dylan Balfour, Lilith Newton, Angie O'Sullivan and Petronella Randell. I would like to thank them for the intense discussion we had at the weekly research seminars, which has led me to develop many ideas in this thesis.

In addition, I would like to thank the incredible St Andrews and Stirling philosophy community for providing an intellectually stimulating and supportive research environment. Besides those already mentioned, I would also like to thank the amazing group of people I met at the SASP programme: Jakob Hinze, Eline Gerritsen, Lorenzo Lazzarini, Saranga Sudarshan, Lara Jost, Katharina Bernhard, Anuj Puri, Matt Clark, Emilia Wilson, Nick Allen, Matteo Nizzardo, Deryn Thomas, Savvas Ioannou, Jose Mestre and many others. Most especially, I would like to thank Jimena Clavel Vázquez, Andrés Villarreal, Paolo Savino, Lixiao Lin and Camille Fouche for sharing this journey with me and for the countless conversations we had about arts and philosophy. They have enriched my life and work. I also want to thank my friends from MLitt: Maddy, Andy, Benjamin, Quan, Harry, Jakob, Silvia, Mo and Stephen. The year we spent together in St Andrews was truly special and beautiful.

I also want to thank the philosophy community at UCL, especially my friends Kay and Androula. Thanks to their excellent taste in music and cinema, a new world has opened to me. I want to thank Luke Fenton-Glynn, my undergraduate thesis supervisor, who said to

me at our last meeting 'this is a work of a philosopher'. Without his encouraging words, I would not have had the confidence to pursue a career in philosophy. I also want to thank my tutor Ziqian Zhou, who taught me how to write philosophy and inspired me to be a teacher just like him.

I had the chance to discuss some ideas from the thesis at the Southern Normativity Group 5th Annual Conference on Normativity, the 6th Stockholm Philosophy Graduate Conference, the Workshop on Intellectual Autonomy at Universidad Autónoma de Madrid and the Pavia University Graduate Conference. I thank the participants for their helpful questions and comments. Part of Chapter 6 contains revised material from my paper 'The Permissible Norm of Truth and "Ought Implies Can"', published in Logos and Episteme 10 (4): 433-440 (2019).

Finally, I would like to thank my old friends: Xiaodan, Qian, Jing, Meng, Zhijian and Jiayuan. I thank my parents, my grandmas, my aunts, and my sisters for their love and support. This thesis is dedicated to the memory of my grandfather, Zuming Wei, who, with my grandma Zehui Zheng, brought me up and always believed in me.

Funding

This work was supported by the John Templeton Foundation PhD Studentship [grant number 58450]; the Aristotelian Society 2020-2021 Student Bursary; and the Royal Institute of Philosophy 2020-2021 Jacobsen Studentship.

Contents

1. Introduction

The Truth Norm of Belief and Reason-based Normative Framework

For belief, the standard of correctness is truth. A belief is correct if and only if it is true.[1] The idea that belief is subject to a norm of truth is widely accepted and central to a variety of philosophical projects across the fields. Some philosophers of mind appeal to the truth norm of belief to individuate belief from other types of mental attitudes. Belief is said to be the type of mental attitude that is subject to a norm of truth. But desire, imagination, hope and other types of mental attitudes are not.[2] Some theorists of truth argue that truth is a substantive property. Truth, as a norm of belief, is a genuine normative property and more than a device of assertion.[3] Some epistemologists employ the truth norm of belief to provide a rationale for truth-conducive accounts of justification. Whether a belief is justified depends on conditions which make it more likely that the belief in question is true.[4] Some theorists of reasons invoke the truth norm of belief to distinguish the 'right kind of reason' from the 'wrong kind of reason' for belief: the right kind of reasons for belief are evidential considerations that bear upon the truth of the belief.[5] The list goes on.

In recent discussions, however, the idea that belief is subject to a norm of truth has come under attack. The anti-normativists challenge the idea that the truth norm of belief is genuinely normative. Some question the authority of the truth norm. Is the truth norm of belief authoritative, and if so, what grounds its authority? Others raise doubts about whether the truth norm can provide any guidance. How can we be guided by the truth norm in forming our beliefs? Still others, argue that the truth norm plays no interesting role in our normative theorizing regarding what to believe since the truth norm of belief does not seem to issue the right epistemic obligations. If so, what sort of claim can the truth norm make on

---

[1] See, e.g., Boghossian (2003); Engel (2013); Gibbard (2005, 2003); Shah & Velleman (2005); Velleman (2000a); Wedgwood (2002, 2007, 2013a, 2013c).
[2] See, e.g., Shah & Velleman (2005); Velleman (2000a); Wedgwood (2002, 2007).
[3] See, e.g., Dummett (1959); Engel (2002); Wright (1992).
[4] On a reliabilist account, a belief is justified if and only if that belief is reliably produced. A reliable belief-forming process is one that produces many more true beliefs than false beliefs (e.g., Goldman 1979, 2008). On an evidentialist account, a belief is justified if and only if that belief fits one's evidence (e.g., Feldman & Conee 1985). A consideration E counts as evidence for a belief p if p is more likely to be true when E obtains than when E does not obtain. For further discussion concerning the truth connection, see, for instance, Conee (1992).
[5] See, e.g., Hieronymi (2005); Reisner (2009).

us regarding what to believe? The normativists have defended the truth norm against such challenges, but they disagree about how to best address these challenges raised by the anti-normativists.

Although contemporary normative theorizing has taken a turn to reason, surprisingly few have approached the debate on a reason-based normative framework in a systematic manner.[6] This thesis aims to advance the debate and meet the anti-normativist challenges by developing and defending a novel account of the truth norm on a reason-based normative framework. In this introductory chapter, I will first lay out the three anti-normativist challenges and explain what I take these challenges to involve on a reason-based normative framework (Sect. 1.1). I will then clarify how I understand the core concepts of this thesis, *belief* and *truth* (Sect. 1.2), and sketch a reason-based normative framework that I will use to defend and develop my account of the truth norm (Sect. 1.3). Finally, I will offer a brief outline of my account of the truth norm and give a sneak preview of later chapters (Sect. 1.4).

1.1 The Truth Norm of Belief: Three Challenges

1.1.1 Why Is the Truth Norm Authoritative?

It is a platitude that belief is subject to a standard of correctness. For belief, correctness is truth: a belief is correct if and only if it is true. Belief is widely held to be subject to a norm of truth, for correctness seems normative:

> '…the concept of correctness seems to be a normative concept. . . The correct belief, if all this is right, seems to be the one [a subject] ought, in this sense, to have.' (Gibbard 2005:338–9)

> '… correctness is clearly a normative property….' (Schroeder 2007:134)

> '… "correctness" is a normative concept, if your belief is incorrect, it has a certain sort of defect – while if your belief is correct, it is wholly free from that defect.' (Wedgwood 2013a:218)

---

[6] It has been widely accepted by philosophers working in normative philosophy that normativity is reason-involving. See, e.g., Parfit (2011); Raz (2010); Schroeder (2007); Skorupski (2007); Dancy (2004); Scanlon (1998, 2014).

A norm, intuitively, is something authoritative or binding for an epistemic agent regarding what to do. But not all standards of correctness are authoritative or binding. Judith Thomson (2008) warns us against the 'tendency to take it that the appropriateness of the word "correct" or "incorrect" in a context is, by itself, a conclusive sign that there is normativity at work in that context' (2008:108). Some standards of correctness may be arbitrary, unjustified or outrightly wrong. In such cases, correctness and normativity can come apart. Suppose you are playing chess, and it is your turn to move. You then learn that if you do not move your knight horizontally, your opponent will be tortured. Are you bound by the rule of chess to not to move your knight horizontally? It seems not. In this case, it does not seem to matter if you move your knight correctly or not – what matters is the safety of the other player. Here is another example. Consider a patriarchal society, in which everything women do is subject to some standard of correctness – from the way they speak to the way they dress. Suppose you are a female member of that society. Are you bound by those standards to live your life? It seems not. Of course, you might be motivated to conform to those standards. But those standards about what is correct or incorrect have no normative authority over what you really may or ought to do. By contrast, moral and rational standards are usually taken to be norms, in the sense that morality and rationality have an authority in making a claim on us regarding what we ought to or may do.

Given this, One challenge to proponents of the truth norm of belief, is to answer what Christine Korsgaard (1996) calls 'the normative question'. The issue is not whether truth is the standard according to which some beliefs count correct, but rather, whether that standard of correctness is normative such that it can get a grip on an agent. Is the truth norm of belief authoritative or binding for an epistemic agent regarding what one ought to or may believe? If so, what grounds its authority?

On a reason-based normative framework, normativity is a matter of normative reasons. If a standard of correctness is genuinely normative, then the fact that an act is correct must be a normative reason counts in favour of performing that act, and the fact that an act is incorrect must be a normative reason for refraining from performing that act. The normative question with respect to the truth norm is best understood in terms of whether there is any normative reason to believe truly (correctly) and not to believe falsely (incorrectly). In other words, the truth norm of belief is authoritative or binding only if the normative reason to believe truly (correctly) and not to believe falsely (incorrectly) is properly grounded.

The first challenge I aim to address in this thesis, then, is to give an account of the normative ground of the reason to believe truly and only truly. Call this the grounding challenge.

1.1.2 Can the Truth Norm Guide?

One widely embraced and frequently invoked idea across normative philosophy is that norms can provide guidance.[7] As Srinivasan nicely puts it for us:

> 'This demand usually issues from an insistence on the "first personal" role of norms, as things that not only provide us with third-personal metrics for assessing others and states of affairs, but also as things that tell us how to think and act – that guide and advise us.' (Srinivasan 2015:279)

So, for norms to play a role in guiding our beliefs and actions, they must be something that can make a difference on the way we think or act from a first-person point of view. Our acceptance of a norm must be able to move us to act or think in accordance with that norm. If the truth norm is genuinely normative, it must be capable of providing guidance in this way.

In recent discussion, the idea that the truth norm can guide has been challenged (e.g., Glüer & Wikforss 2009, 2010, 2013). On a standard inferentialist model of guidance, to be guided by a norm of the form 'φ if and only if C', is for one to first form a belief about C and then φ (or not φ) on the basis of the belief about C. But it seems impossible to be guided by the truth norm on the inferentialist model. For to form a belief about C, in this case, is to form a belief about whether p is true. But to form a belief about whether p is true is just to form a belief about p. So, it seems that to be guided by the truth norm to form a belief about p, one already has to form a belief about p, which is the very same thing one seeks guidance for in the first place. Or so the objection goes.

Instead of trying to account for how the truth norm can guide on the inferentialist model, I will argue that the inferentialist model is not the only way to understand guidance. I will propose an alternative, reason-responsive model of guidance, according to which to be guided by a norm is to respond to the relevant normative reasons. On a reason-based normative framework, the challenge about guidance is best understood in terms of *responding*

---

[7] See, e.g., Williams (1981); Nagel (1989); Hursthouse (1999); Cruz & Pollock (1999); Thomson (2008); Darwall (1983); Gibbard (1990); Korsgaard (1996).

to a reason. The truth norm of belief cannot provide guidance if the reasons to believe truly and only truly are not reasons that we can respond to. If the truth of p is a normative reason for belief, it must be possible for one to believe p and to refrain from believing not-p in response to the reason that a p is true.

The second challenge I aim to address in this thesis, then, is to offer an account of what it takes to believe in response to truth and show that it is possible to believe p and to refrain from believing not-p for the reason that a p is true. Call this the guidance challenge.

1.1.3 Can the Truth Norm Make Any Claim Regarding What to Believe?

Another compelling idea about norms is that norms can make a claim on us regarding what to do. Norms tell us under what circumstances what sort of responses are required, permitted or forbidden.[8] Following von Wright (1963), we can distinguish between norms of obligation and norms of permissibility:

> 'The character of a norm depends upon whether the norm is to the effect that something ought to or may or must not be or be done… Norms of the "ought" – character can also be called obligation-norms, and norms of the "may" – character permissive norms.' (von Wright 1963:71)

For example, the norm of promise-keeping is usually taken to be an obligatory norm. If you make a promise, you ought to keep it. The norm of evidence is an example of a permissible norm.[9] If you have sufficient evidence for p, you may believe p. These norms can make a claim on us in the sense that we can be held accountable for violating the norms and any violation without good excuses would leave us open to blame.

Recently, it has been argued that the truth norm plays no interesting role in our normative theorizing regarding what to believe. The truth norm of belief seems to issue excessive, unsatisfiable obligations (e.g., Bykvist & Hattiangadi 2013; Glüer & Wikforss 2013; Hattiangadi 2010). This motivates a permissible account of the truth norm, according to which one is permitted to believe p if and only if p is true (e.g., Whiting 2010, 2012, 2013). However, as we shall see, there are a variety of cases in which the epistemic agent is arguably

---

[8] I follow most normativists in this debate and understand normativity in prescriptive terms, which is closely related to the idea that norms provide guidance. For alternative construal of normativity in evaluative and teleological terms, see, McHugh & Whiting (2014).
[9] Pollock (1987) goes further and claims that all epistemic norms are permissible norms.

neither obliged nor permitted to believe truly and only truly. If so, what sort of claim can the truth norm make on us?

On a reason-based normative framework, the deontic concepts 'ought' and 'may' are understood in terms of the weight of reasons that stand in a relation to an agent. If a reason R favours φ-ing has a requiring weight for an agent S, then S ought to φ when R obtains. If R has a permitting weight for S, then S may φ when R obtains.

The third challenge I aim to address in this thesis, then, is to offer an account of the weight of the reason to believe truly and only truly, such that it can accommodate a wide range of intuitive cases but also allows us to make good of the claim that the truth norm can play an interesting role in our normative theorizing regarding what we ought to or may believe. Call this the weighting challenge.

Before proceeding to outline my answers to these challenges, let me first say a few words about how I understand the core concepts of this thesis, *belief* and *truth*, and sketch the reason-based framework on which I will defend and develop the truth norm of belief.

1.2 Three Preliminaries

1.2.1 What is Belief?

Most philosophers take belief to be a type of mental attitude held towards propositions. By propositions I mean essentially truth-evaluable abstract entities.[10] Some philosophers also hold that a belief comes in degree, that is, a belief as an attitude comes in various 'degrees of confidence' about the truth of the proposition involved. The truth norm of belief I defend and develop in this thesis concerns 'flat-out' belief and I will simply assume that a belief comes with a high degree of confidence.[11] The exact characterization of the attitude-type of what we call 'belief' is a matter of substantial debate. There are three main accounts of belief in the literature: *representationism*, *functionalism*, and *dispositionaliam.*[12]

Very briefly, representationists hold that to believe p is to stand in a particular relation to an internal mental representation with the content 'p' stored in a 'Belief Box' (to use a

---

[10] Some take propositions to be unstructured sets of centred possible worlds. Some take propositions to be structured entities whose constituents are mind-independent abstract concepts or Fregean senses. Others take propositions to be structured entities comprising Russellian concrete objects, properties and relations. I assume that propositions are structured and Fregean.

[11] For an overview of various issues concerning the relation between degrees of belief and outright belief, see, e.g., Ebert and Smith (2012). Much of our discussion will also apply to the parallel probabilistic norm governing one's degree of belief or credence about p.

[12] For a review, see, e.g., Schwitzgebel (2006).

familiar metaphor). It is poised to perform some specific set of causal roles such as causing certain behaviour in the presence of appropriate desires.[13] Functionalists hold that mental states are functional states. A functional state is a state that plays a type of causal role in a system. On a functionalist account, a belief p is defined by its causal relations to both what tends to cause it (stimuli and sensory inputs such as perception) and what it tends to cause (behaviour such as asserting p or other mental states of the system).[14] Dispositionalism is the view that to believe p is nothing more than having a cluster of dispositions, which may include behavioural dispositions as well as 'cognitive' and 'phenomenal' dispositions (e.g., Schwitzgebel 2002, 2013).[15]

It is fairly uncontroversial to say that these three accounts of belief overlap, in particular with the more recent developments. On all three approaches, belief is the type of propositional attitudes which involves dispositions to behave (mentally and bodily) as if the proposition it stands in relation to were true. This minimal conception of belief is all I will need to develop my account of the truth norm of belief. I take this to be one of the advantages my account enjoys over a rival, constitutive account of the truth norm, which relies on the claim that belief is by nature subject to a norm of truth (Sect. 2.3). By not presupposing a substantive theory of belief, my account of the truth norm allows us to import the result of our investigation to shed light on the normative nature of belief.

1.2.2 What is Truth?

The truth norm of belief claims that a belief is correct if and only if is true. But what is it for a proposition to be true? There are three broad families of views regarding the nature of truth.

Very briefly, on a monist conception of truth, there is only one way of being true: truth is understood as correspondence to facts or states of affairs that obtain (correspondence theories of truth); as coherence with some specific set of propositions (coherence theories of truth); as satisfaction of some epistemic conditions (epistemic theories of truth); as satisfaction of some utilitarian conditions (pragmatist theories of truth). On a pluralist conception of truth, there is more than one way of being true across different truth-apt discourses. On a deflationist conception of truth, truth is not a genuine property. The theory

---

[13] See, e.g., Fodor (1968, 1987, 1990); Dretske (1988); Nichols and Stich (2003); Quilty-Dunn & Mandelbaum (2018a).
[14] See, e.g., Armstrong (1973); Loar (1981).
[15] See, e.g., Ryle (1949); Sellars (1956); Davidson (1984); Lewis (1972).

of truth is confined to a simple principle, the equivalence schema: it is true that p if and only if p.[16]

I am taking it for granted that true propositions are true in virtue of the way that the world is. Following Armstrong (2004), I assume that every true proposition has a truthmaker.[17] I take truthmakers, among other things, to be facts. So, the proposition that Tolstoy is a Russian writer is true just in case there is some fact that makes it true, e.g., the fact that Tolstoy is a Russian writer.[18] The theory of truthmakers is not a theory of truth – it does not aim at providing a definition of truth or at providing an account of the nature of truth. Hence, by merely appealing to a theory of truthmaker, the truth norm of belief defended in this thesis remains neutral regarding various theories of truth.

Although the truthmaker theory is in the vicinity to correspondence theories of truth and shares a realist intuition about truth, the arguments made in developing my account of the truth norm are independent of any particular theory of truth. I take this to be one of the advantages my account enjoys over a rival, value-based account of the truth norm, which is motivated by the idea that truth is a value (Sect. 2.2). Again, by not presupposing a substantive theory of truth, we can draw on the result of our investigation to shed light on the normative nature of truth.

1.2.3 What is Reason?

1. Reasons and Grounds. Contemporary normative theorizing has taken a turn to reason. Normativity is a matter of normative reasons. Normative reasons are facts that count in favour of various responses.[19] By responses I refer to things that are responsive to reasons, such as actions, intentions and attitudes.[20]

---

[16] For an overview of various theories of truth, see, e.g., Engel (2002).

[17] For further discussion on truthmaker theory, see, e.g., Beebe & Dodd (2005)

[18] There are true propositions made true by more than one fact and every fact makes many true propositions true. For example, the fact that Tolstoy is a Russian writer also makes an infinite number of disjunctive propositions true: propositions that take 'Tolstoy is a Russian writer' as one of its disjunct. The proposition that Tolstoy is a Russian writer is also made true by facts about Tolstoy that imply that he is a Russian writer.

[19] In this thesis I will take it for granted that *reasons against* a response are *reasons for* refraining from adopting that response (cf., Nagel 1970:47; Schroeder 2007: Chp.7). I will leave aside the ongoing debate concerning the precise logical relationship between reasons against and reasons for. For further discussion on this issue, see, e.g., Snedegar (2018).

[20] Parfit (2001:21-5) introduces a further distinction between state-given reasons and object-given reasons for a propositional attitude. Piller (2006) labels the distinction as one between 'content-related' and 'attitude-related' reasons. The idea is roughly this. A reason for an attitude is *object-given* if the fact in favour of that attitude bears upon the attitude's object, e.g., the proposition in question,

Here are some intuitive examples of reasons. The fact that there is a climate emergency is a reason to cut our carbon footprints. The fact that animals suffer is a reason to refrain from eating meat. The fact that the upcoming Edinburgh Book Festival has an excellent programme is a reason to intend to attend the festival. The fact that Brewdog is crowded is a reason to refrain from intending going there. The fact that the pavement is wet is a reason to believe that it has rained. The fact that Smith has an alibi is a reason to refrain from believing that Smith committed the crime.

A list of examples, of course, does not tell us what it takes for some fact to count in favour of a response. It is *not* my present project, however, to figure out the metaphysics of normative reasons. I will not defend any particular view about what it takes for a fact to be a normative reason.[21] Rather, I am taking it for granted it that a reason can be grounded in a variety of desire-based, value-based, constitutive or practice-based facts. I will assume that for a fact R, a response φ, R constitutes a reason in favour of φ-ing if one of the following grounding facts obtains:

(Desire-based)	If R were to obtain, φ-ing would promote desired states of affairs.
(Value-based)	If R were to obtain, φ-ing would promote valuable states of affairs.
(Constitutive)	If R were to obtain, it is constitutive of a certain act/attitude that R would count in favour of φ-ing.

whereas a reason for an attitude is *state-given* if the fact in favour of that attitude bears on having that attitude itself, regardless of its content. There is an ongoing debate in the literature about whether state-given reasons for belief exist. Several philosophers hold that there are only object-given reasons for belief. They claim that people who think there are state-given reasons for belief confuse the reason for having a belief with reason for bringing about that belief (See, e.g., Parfit 2011, Appendix.A; Skorupski 2010:54-5 and Way 2012). The fact that you will save one million lives may give you a reason to *bring about* the belief that the moon is made of blue cheese. There are many ways you might try to manipulate yourself into believing it. But the fact that you will save one million lives is not a reason for having that belief. I am inclined to think that this distinction between object-given and state-given reasons for belief does not really hold water. We cannot talk about a belief attitude in abstract of its content and we cannot talk about a belief attitude in abstract of its being an attitude playing belief-like causal roles in our mental economy, for a belief p is a *propositional attitude.* In this thesis, I make no such distinction between object-given and state-given reasons and none of the arguments I make will rely on this distinction.

[21] Very briefly, reason fundamentalists hold that at least some facts about reasons are normatively fundamental, in the sense that they are not fully grounded in other normative facts (e.g., Parfit 2011; Raz 2010; Scanlon 1998, 2014). Others think facts about reasons are ultimately grounded in facts about desires (e.g., Smith 1994; Schroeder 2007); in facts about values (e.g., Finlay 2006; Maguire 2016); or in facts about rationality (e.g., Korsgaard 1996).

(Practice-based) If R were to obtain, φ-ing would constitute a justified social practice.

We can use this schema as heuristics to determine whether R is a properly grounded normative reason for φ-ing. Here are some examples. Life with dignity constitutes desired and valuable states of affairs. The fact that refugees need food and shelter is a desire-based and value-based reason to help them. It is constitutive of the act of promising that making a promise counts in favour of keeping that promise. The fact that one made a promise is a constitutive reason to keep that promise. Punishing criminals in accordance with the law constitutes a justified social practice. The fact that someone commits a crime is a practice-based reason to punish that person in accordance with the law.

The question we will be investigating in Part I of the thesis is whether the reason to believe truly and only truly can be properly grounded. And if so, whether that reason a desire-based, value-based, constitutive or a practice-based reason. I will defend a novel account of the truth norm, according to which, the reason to believe truly and only truly is practice-based.

2.Respondign to Reasons. Normative reasons should be distinguished from *motivating reasons.*[22] Normative reasons, as we have said, are facts that count in favour of φ-ing. The fact that a project will lift one hundred families from poverty is a normative reason to invest in that project. Normative reasons can stand in a relation with agents in a given situation. The fact that a project will lift one hundred families from poverty is a reason for, say, the local mayor, to invest in that project with public fund. Motivating reasons are considerations for which an agent acts, which may or may not be facts that count in favour of that act.[23] For example, the reason for which Sam got drunk at a party was peer pressure to drink more. But the fact that there was peer pressure to drink more does not count in favour of getting drunk at a party.

---

[22] Normativity is a matter of normative reasons, not motivating reasons. Unless I state otherwise, in the remainder of this thesis, 'reason' refers to normative reason. For a helpful overview of the distinction between normative reasons and motivating reasons for belief, see, e.g., Sylvan (2016a, 2016b).

[23] There is an ongoing debate about the ontology of motivating reasons: whether they are facts, mental states, or something else. I am taking it for granted that at least some motivating reasons are facts, i.e., normative reasons. For a recent discussion, see, e.g., Alvarez (2018).

Not every motivating reason is a normative reason, but it must be possible for a normative reason to be a motivating reason when an agent comes to stand in a relation to that normative reason. It is an earmark of normative reasons that they are the sort of things that one can respond to. As Bernard Williams puts it, 'If there are reasons for action, it must be that people sometimes act for those reasons' (1981:102).

The question we will be investigating in Part II of the thesis concerns whether the truth of a proposition can be a reason that one can respond to. I will defend a novel view of what it takes to respond to the reason to believe truly and only truly.

3. Reasons and Weights. When reasons stand in a relation to an agent, they have *weights* and can be weighed against one another. The reason to drive carefully to avoid an accident is more demanding on an agent than the reason to turn on the radio to enjoy music. These reasons have different weights and they weighed against each other in making a claim on an agent regarding what to do, i.e., whether to focus on driving or to turn on the radio.

Some reasons are weighty enough to issue a permission or even a requirement on an agent regarding what to do. If a reason R favours φ-ing has a *requiring weight* for an agent S, then S ought to φ when R obtains. If R has a *permitting weight* for S to φ, then S may φ when R obtains. Here are some examples. The fact that my uncle is allergic to nuts seems to be a requiring reason for him to refrain from eating nuts. He ought to refrain from eating nuts since he is allergic to nuts. The fact that there is a climate emergency seems to be a permitting reason for me to engage in civil disobedience, but it falls short of having a requiring weight. I am permitted to engage in civil disobedience since there is a climate emergency though I do not have an obligation to do so.

As we will see in chapter 7, the distinction between requiring and permitting weight is more complicated than what has been illustrated by the above examples. Weights of reasons are holistic and context-sensitive. A reason to φ can, in one context have a certain weight for an agent, and in another context a different one. The weight of a reason is determined by what else is true in a given situation, or so I shall argue.

In Part III of the thesis, I will defend a novel, variantist account of the truth norm, according to which the weight of the reason to believe truly and only truly can vary from case to case, depending on the circumstances.

1.3 The Plan

The remainder of the thesis is divided into three parts. Part I concerns the authority of the truth norm and aims to meet the grounding challenge. Chapter 2 considers and criticizes two popular views: the evaluative account and the constitutive account of the truth norm. I argue that the existing attempts to ground the truth norm by appealing to desire-based, value-based and constitutive reasons fail. Chapter 3 develops a practice-based account of the truth norm, according to which the reason to believe truly and only truly is grounded in what I call the *T-practice*, a justified social practice governed by the truth norm of belief.

Part II concerns the guiding capacity of the truth norm and aims to meet the guidance challenge. Chapter 4 discusses difficulties facing the standard inferentialist account of how the truth norm can guide, in light of emerging empirical evidence about our system of belief-formation. I propose an alternative, reason-responsive model of epistemic guidance. On this view, to be guided by the truth norm is to respond to the T-practice-based reason. In Chapter 5, I develop and defend a novel hybrid dispositionalist/phenomenological account of what it takes to respond to the T-practice-based reason. I argue that believing something for the T-practice-based reason is a kind of mental action with distinctive phenomenology. For one to respond to the T-practice-based is for one to either manifest the relevant disposition to respond to the T-practice-based reason or manifest a distinctive phenomenology of the mental action of believing something for the T-practice-based reason.

Part III concerns the normative force of the truth norm and aims to meet the weighting challenge. Chapter 6 criticizes two leading invariantist views, according to which the T-practice-based reason to believe truly and only truly is a requiring/permitting reason across all contexts. I argue that both views are vulnerable to counterexamples. Chapter 7 develops and defends a variantist account of the truth norm which says that whether the T-practice-based reason is a requiring or permitting reason for belief can vary from case to case, depending on the circumstances.

Part I. Why Is the Truth Norm of Belief Authoritative?

Preview of Part I

In the previous chapter, I laid out three anti-normativist challenges to the truth norm of belief in recent literature: (1) the grounding challenge; (2) the guidance challenge; and (3) the weighting challenge. The aim of my thesis is to offer novel responses to the three challenges and defend the truth norm of belief on a reason-based normative framework. I explained what I take these challenges to involve on a reason-based normative framework and introduced the main moving parts.

In Part I, I turn to addressing the grounding challenge. On the reason-based framework, if the truth norm of belief is authoritative or binding, then the reason to believe truly (correctly) and to refrain from believing falsely (incorrectly) must be properly grounded. So, what, if anything, grounds the reason to believe truly and only truly? In chapter 2, I will consider and criticize existing attempts to ground the reason on value-based, desire-based and constitutive reason grounding facts. In chapter 3, I will defend a practice-based account of the truth norm according to which the reason to believe truly and only truly is grounded in a justified social practice.

2. The Grounding Challenge

On the reason-based normative framework, normativity is a matter of reasons. Reasons are facts that count in favour of various responses. Whether the truth norm of belief is authoritative, I suggested, is a question about whether the reason to believe truly and only truly can be properly grounded. Recall that, for a fact R, a response φ, R constitutes a reason in favour of φ-ing if one of the following reason grounding facts obtains:

(Desire-based)	If R were to obtain, φ-ing would promote desired states of affairs.
(Value-based)	If R were to obtain, φ-ing would promote valuable states of affairs.
(Constitutive)	If R were to obtain, it is constitutive of a certain act/attitude that R would count in favour of φ-ing.
(Practice-based)	If R were to obtain, φ-ing would constitute a justified social practice.

In this chapter, I will first provide a brief overview of two main views about the authority of the truth norm in the current literature: the evaluative and the constitutive account of the truth norm. I will identify and explain what I take to be a fundamental challenge to both views: to show that the relevant reason grounding facts obtain (Sect. 2.1). I will then argue that neither the evaluative account nor the constitutive account of the truth norm has successfully shown that their preferred reason grounding facts obtain (Sect. 2.2-2.3).

2.1 The State of the Debate

2.1.1 The Evaluative Account of the Truth Norm

There are two prominent views about the authority of the truth norm in the current literature. One is the evaluative account of the truth norm. According to this view, the truth norm of belief is authoritative because true beliefs are good (e.g., Alston 1985; Bonjour 1985; Goldman 1999; Sosa 2003; Lynch 2004a, 2004b, 2009; Fassio 2011; McHugh 2012).

The evaluative account of the truth norm has its historical roots in the American pragmatist tradition.[1] At the heart of the pragmatist programme is the thought that true ideas are those that can be verified, and the verification-processes involve putting one's ideas in practice to see whether they *pay off*:

> 'Pragmatism...asks its usual question. "Grant an idea or belief to be true," it says, "what concrete difference will its being true make in any one's actual life? How will the truth be realized? What experiences will be different from those which would obtain if the belief were false? What, in short, is the cash-value in experiential terms?' (James 1907:142)

> 'True ideas lead us into useful verbal and conceptual quarters as well as directly up to useful sensible termini. They lead to consistency, stability and flowing human intercourse. They lead away from eccentricity and isolation, from foiled and barren thinking.' (James 1907:147)

On a Jamesian pragmatist conception of truth, true beliefs are *identified with* beliefs that are valuable. Most philosophers, of course, reject Jamesian pragmatist conception of truth for obvious reasons.[2,3] But there is something intuitive about Jamesian pragmatist conception of truth that many philosophers are still attracted to, namely, the idea that truth is something

---

[1] Notice that, as Haack (1976) observes, other historical proponents of pragmatist conceptions of truth, such as Schiller, Dewey and Peirce hold views that are in fact more appropriately characterized as epistemic conceptions of truth.

[2] For one thing, there are clearly true and false propositions whether we believe in them or not. For another, there are clearly true beliefs that are not useful and false beliefs that are. Moreover, the pragmatist theory of truth seems to imply that truth is subjective since what is useful depends on our desires and goals and the things we regard as useful do not necessarily converge. What is useful for you may not be useful for me and hence what is true for you may not be true for me. But this is not how we usually understand truth. To make such position more plausible, it seems that one has to qualify and restrict what counts as useful belief, a move which most likely would involve importing some elements from correspondence, coherence, or epistemic theories of truth. If so, it is no longer clear that Jamesian pragmatic conception of truth is a distinctive conception of truth from which we can derive the claim that true beliefs are valuable (see, e.g., Engel 2002; Russell 1940; Schmitt 1995).

[3] Interestingly, some commentators point out that Jamesian pragmatist conception of truth is perhaps best understood as offering a criterion for *true belief*, rather than as a definition for truth. For example, Engel (2002:36) suggests that a pragmatic conception of true beliefs may be understood as a theory which identifies the truth conditions of belief with their utility conditions. One contemporary development of this line of thought is teleosemantics (e.g., Millikan 1984). It explains the *mental content* of a belief in terms of that belief's biological function, which is said to lie in the way that belief generally leads to successful actions in the long run.

valuable or desirable and that it plays an important role in regulating our mental economy and guiding our daily life. The evaluative account seeks to meet the grounding challenge by appealing to this intuitive idea that true belief is good. If believing p and refraining from believing not-p would promote valuable or desired states of affairs when p were true, then the value-based or desire-based grounding fact would explain why the truth of p is a reason to believe p and to refrain from believing not-p.

Despite its intuitive appeal, the evaluative account of the truth norm faces a familiar problem: the evaluative account of the truth norm seems to offer the wrong kind of explanation of why truth is a reason for belief. To begin with, it seems that the truth of p is a reason to believe p and to refrain from believing not-p regardless of whether that belief is desirable/valuable. Even if believing p is *not* desirable/valuable, it seems that the truth of p is nevertheless a reason counts in favour of believing p. Moreover, there are many ways in which a belief might be desirable/valuable that are independent from whether the belief is true or false. For example, believing that the pandemic has ended might make me happy and promote valuable and desired states of affairs. But the fact that a belief makes me happy is the wrong kind of reason for believing p. If one is committed to the view that the reason to believe p and to refrain from believing not-p is value-based or desire-based, one seems to be forced to accept other value-based or desire-based reasons for belief such as the fact that believing so and so makes one happy. But this seems wrong.

Some philosophers, of course, endorse what is known as the pragmatist view about epistemic reasons, according to which, there can be value-based or desire-based reasons for belief which do not bear upon the truth of a belief (e.g., Papineau 2013; Rinard 2015, 2017, 2018, 2019; Reisner 2009; Steglich-Petersen & Skipper 2019; Woods & Maguire 2020). On this view, the fact that believing p promotes some valuable or desired states of affairs could be a reason for believing p even if p is not true. After all, desire-based and value-based reasons are common in the practical realm. Rinard (2017, 2018, 2019), for instance, thinks it is a distinctive virtue of pragmatism that it unifies epistemic normativity and practical normativity: they are not two distinct realms, but rather, they are on the same normative landscape. As such, the wrong-kind-of-reason problem does not bite. Perhaps so. But pragmatism about epistemic reasons is not an insignificant commitment.

Regardless of the outcome of the debate concerning pragmatism about epistemic reasons, the concern I have for the evaluative account is more fundamental. For the evaluative account to successfully meet the grounding challenge, its proponent must show that the relevant value-based or desired-based grounding facts obtain. This, I submit, is

where the evaluative account falls short of. In Sect. 2.2, I will spell out three possible ways to support the claim that the reason to believe p and to refrain from believing not-p when p is true is value-based/desired-based. But we shall see that none of the arguments is satisfactory.

2.1.2 The Constitutive Account of the Truth Norm

Another prominent view about the authority of the truth norm of belief in the current literature is the constitutive account of the truth norm. According to this account, the truth norm of belief is authoritative because it is *constitutive* of belief that it is subject to a norm of truth (e.g., Korsgaard 1996; Wedgwood 2007; Boghossian 2008; Shah 2003; Gibbard 2005; Velleman 2000a; Shah and Velleman 2005).

The idea that there is a conceptual connection between truth and belief is deeply rooted in the analytic tradition. Frege defines judgment (occurrent belief) as 'not the mere comprehension of a thought but the admission of its truth' (1948:216 fn7).[4] Bernard Williams (1973) famously labels belief as the kind of attitude that aims at truth:

> 'Beliefs aim at truth…Truth and falsehood are a dimension of an assessment of beliefs as opposed to many other psychological states or dispositions… To believe that so and so is one and the same as to believe that that thing is true.' (Williams 1973:136-137)

More recently, Velleman (2000b) claims that it is a *conceptual truth* about belief that belief is subject to a norm of truth:

> 'The concept of belief just is the concept of an attitude for which there is such a thing as correctness or incorrectness, consisting in truth or falsity. For a propositional attitude to be a belief just is, in part, for it to be capable of going right or wrong by being true or false.' (Velleman 2000b:16)

[4] This characterization is found throughout Frege's work, see, for instance, Basic Laws of Arithmetic (2013: §5); 'Introduction to Logic' (1991:185); 'A Brief Survey of my Logical Doctrines' (1991:198); 'My Basic Logical Insights' (1991:251); 'The Thought' (1956:329); 'Notes for Ludwig Darmstaedter' (1991:253); and 'Sources of Knowledge in Mathematics and the Mathematical Natural Sciences' (1991:267).

Indeed, within the analytic tradition, the idea that the truth norm of belief is constitutive of belief has been widely accepted.[5] The constitutive account seeks to meet the grounding challenge by appealing to this intuitive idea that truth is the constitutive norm of belief. If it is constitutive of belief that if p were true, p would count in favour of believing p and refraining from believing not-p, then the constitutive grounding fact would explain why the truth of p is a reason to believe p and to refrain from believing not-p.

One might find the constitutive account attractive for it avoids the wrong-kind-of-reason problem. The constitutive reason to believe p and to refrain from believing not-p is distinctive of the attitude type of belief. It does not force one to accept other desire-based or value-based reasons for belief. But despite its strong appeal, the constitutive account of the truth norm also faces a familiar problem: the constitutive account of the truth norm seems to have limited 'authority'. Consider the constitutive rules of chess. As far as you are playing the game of chess, there is a reason against moving your knight horizontally. But you have no reason to refrain from moving your knight horizontally if you are not playing a game of chess. Whether the rules of chess are normatively binding depends on whether one is playing the game of chess. Likewise, as far as you are playing the game of belief, the truth of p is a reason for you to believe p and to refrain from believing not-p. But you have no reason to do so, if you are not playing the game of belief with respect to p. Whether the truth norm of belief is normatively binding depends on whether one is playing the game of belief. So, the constitutive grounding fact could explain why the truth of p is a reason to believe p and to refrain from believing not-p, only if one were to form a *belief* about p. But this seems wrong. The fact that 2+2=4 is true is a reason for believing that 2+2=4, regardless of whether one forms a belief about that proposition.

Perhaps it is all right to just bite the bullet. After all, unlike the game of chess, the game of belief is not something we can easily opt out (e.g., Feldman 2000). The constitutive grounding fact could explain why the truth of p is a reason to believe p and to refrain from believing not-p for ordinary epistemic agents since playing the game of belief is just part of life.

---

[5] I follow Burge (2010:58) in understanding constitutive conditions as conditions under which something is what it is. It is a condition that provides a fundamental explanation of a thing being that kind of thing. For example, being H_2O is constitutive of water. It is in virtue of having that chemical composition that a substance is water. The chemical composition of H_2O also provides a fundamental explanation and illuminates the nature of water. Here is another example. The rules of chess are constitutive of chess since they ground our understanding of what it is to play chess and it is in virtue of the fact a game is governed by chess rules that a game is chess.

Regardless of whether one finds this response satisfying, the concern I want to raise in this chapter is more fundamental. For the constitutive account to successfully meet the grounding challenge, its proponent must show that the relevant constitutive grounding fact obtains. This, I submit, is where the constitutive account falls short of. In Sect. 2.3, I will spell out three possible ways to support the claim that the reason to believe p and to refrain from believing not-p when p is true is constitutive of belief. But we shall see that none of the arguments is satisfactory.

2.2 Why Truth is not a Value-based/Desire-based Reason for Belief

2.2.1 True Belief and Utility

Let me begin with the intuitive idea that true belief is useful. From a practical standpoint, true beliefs are useful for getting what we want, after all, true and only true beliefs map the world in such a way that they enable us to navigate through it successfully. Without a true belief about the bus timetable, for example, I would be late for work. We seem to have a straightforward argument for the obtaining of the desire-based/value-based grounding fact: for any p, believing p and refraining from believing not-p when p were true would promote desired/valuable states of affairs because true beliefs are useful in getting what we want.

The problem with this argument is that true beliefs do not always pay off – although true belief is, *in general* useful for getting what we want, it is unlikely that *every* true belief would promote desired states of affairs or be instrumentally valuable in this way. There is a myriad of true propositions that if one were to believe them, would not produce any desired/valuable outcomes and may even produce undesired/disvaluable outcomes.

One initial response might suggest that no belief that has the property of being true is totally useless because it might later turn out to be useful. That is, every true belief has a potential to be of value when the right circumstances arise. Although my beliefs about trivial facts may not have any obvious use, when the right circumstances arise, they might become valuable, say, for correctly answering some odd questions in a pub quiz. But, importantly, to say that every true belief has a *potential* instrumental value is not the same as saying that every true belief has an instrumental value. The fact that believing p and refraining from believing not-p when p were true would *potentially* promote desired/valuable states of affairs cannot ground the reason to believe p and to refrain from believing not-p when p is true. The truth

of p constitutes a value-based/desire-based reason only if believing p and refraining from believing not-p when p were true *would promote* desired/valuable states of affairs.

A more plausible interpretation of the idea behind the initial response is to say that, in the long run, believing truly and only truly even with respect to trivial propositions is beneficial. The thought is familiar. Following rules of traffic in a sparsely populated rural town may not have any immediate beneficial effect, but in the long run, following rules of traffic pays off. Likewise, although true beliefs do not always pay immediately, believing truly and only truly eventually pays off.

I am very much in agreement with this general line of thought. However, by adopting this line of response, one accepts the idea that the locus of instrumental value is not an individual belief or an act of believing, but the manifestation of a pattern of doxastic attitudes, or, as I would prefer to call it, a *practice* of belief management governed by the truth norm of belief. The idea that, in the long run, the practice of believing truly and only truly even with respect to trivial propositions is beneficial to both individual and society at large in fact motivates a practice-based account of the truth norm which I will defend in the next chapter. Therefore, I do not think proponents of the evaluative account can avail themselves to this response, without committing themselves to a significant departure from their original view.

In short, since it is not the case that every true belief is instrumentally valuable, the argument from utility for the evaluative account fails to hold water. But in any case, the utility of true belief is not what draws most philosophers to the evaluative account of the truth norm. Most philosophers regard true belief as valuable/desired for its own sake. This leads us to the next two arguments for the evaluative account.

2.2.2 True Belief and Intrinsic Value

It is a popular idea among philosophers that the value of true belief does not depend on anything other than the fact that it has the property of being true (e.g., Zagzebski 2003; Kvanvig 2003; Lynch 2004a, 2004b). True belief is valuable for its own sake.[6] Although true belief might lack instrumental value (for the existence of such value depends on external factors), it necessarily has intrinsic value in virtue of possessing the property of being true. Here we have another straightforward argument for the obtaining of the value-based

[6] Notice that some philosophers distinguish between intrinsic value and final/non-instrumental value (e.g., McGrath 2005; Korsgaard 1983; Langton 2007). Roughly, something is intrinsically valuable if it is valuable in virtue of its intrinsic properties and something is finally/non-instrumentally valuable, if it valuable for its own sake. I do not make the distinction here.

grounding fact: for any p, if p were true, believing p and refraining from believing not-p would promote valuable states of affairs because true belief is intrinsically valuable.

But why should we accept the premise that true belief has intrinsic value? Perhaps you think there is not much one can say or need to say since, as Wittgenstein puts it, it is where our spade is turned. It is difficult to justify why something is intrinsically valuable, for there is nothing one can easily appeal to other than the nature of that thing itself. But if one appeals to the nature of belief to explain why true belief is valuable, the evaluative account seems to collapse into a constitutive account of the truth norm.

Interestingly, Kvanvig (2003:154) thinks that the difficulty in explaining the value of true belief itself indicates that the value of true belief is intrinsic:

> 'When we find something of intrinsic value and attempt to explain its value, we are forced to resort to explanations that are apparently circular. For example, many take pain to be disvaluable. When questioned concerning its disvalue, it is hard to cite any explanation of its disvalue other than the fact that pain hurts. This is no explanation, however. Instead, what it indicates is an inability to explain the disvalue of pain in any other terms, indicating an account of the disvalue of pain on which pain is disvalued intrinsically. The same should be said about true belief.' (Kvanvig 2003:154)

According to Kvanvig, we cannot give any explanation of the intrinsic value of true belief, other than appealing to the fact that it is true. The lack of non-circular explanation for the intrinsic value of true belief is a reason to think that true belief is intrinsically valuable.

But this argument is unpersuasive for two reasons. For one thing, a lack of non-circular explanation of the intrinsic value of true belief may very well indicate that true belief in fact has no value as opposed to indicating that it has intrinsic value. Isn't it more intuitive to think that the lack of an explanation of why my belief about the number of leaves on a tree has value indicates that my belief has no value as opposed to that it has intrinsic value?

Furthermore, we can think of many examples of non-circular explanation of intrinsic value. For instance, Bach's Goldberg Variations is presumably intrinsically valuable – assuming that it is a piece of art that we value for its own sake, not merely for the pleasure deriving from listening to it. Yet, this does not mean that we cannot offer any explanation as to why it is so. In fact, we can explain its value by analysing, among other things, its musical

language.[7] There is no obvious reason why, if true belief has intrinsic value, we cannot offer a non-circular explanation of its value.

In contrast to Kvanvig's suggestion, one plausible indicator of intrinsic value, I think, is the way we are disposed to react to a putative value-bearer. If a true belief is intrinsically valuable, then at least some agent will value that belief for its own sake, which, plausibly, involves some dispositions to act as if that belief is valuable, for example, by expressing one's positive opinion of that true belief, among other things. But do we value true beliefs that are uninteresting or trivial to us? Are we disposed to react to them as if they are valuable? It seems not.

Here is one way to spell out this worry. Heal (1988) invites us to imagine a person with a notebook setting out to learn the number of cars parked in different streets simply because there are such truths to be gathered about the car numbers. If true beliefs about car numbers are valuable for their own sake, then this person's action should be intelligible to us. But Heal thinks we cannot make sense of this person's action.

Of course, one might imagine that this person has nothing better to do with his time than going around to learn car numbers. But in that case, there are more plausible interpretations of his action than saying that his action manifests some disposition to value true beliefs about car numbers for their own sake.

This case suggests that we are not disposed to react to uninteresting truths as if they are valuable. The fact that we do not value such truths is an indicator that they lack value. It is thus far from clear that the claim that true belief has intrinsic value is a truism in which our spade must turn. The intrinsic value of true belief seems nothing more than a myth.

But perhaps there is another way to explain the unintelligibility of the car numbers case. The reason that person does appear unintelligible to us is maybe because there should be other outweighing concerns preventing him from valuing those true beliefs about car numbers. The thought is, although true beliefs about car number have intrinsic value, that value is negligible and easily outweighed by other concerns. But this apparent unintelligibility is simply a reflection of our limitation, not a rebuttal of the claim that true belief has intrinsic value. Surely, there is nothing unintelligible about an ideal epistemic agent (say, someone with unlimited cognitive resources living in a cost-free environment) manifesting dispositions to value every true belief for its own sake. Lynch puts the point as follows:

---

[7] See also, e.g., Hazlett (2013:32); Horwich (1998:63); and DePaul (2009:132).

> 'There are all sorts of trivial truths that are not worth believing, given my limited intellect and time. Nonetheless, were these limits not in place – were it to be the case that believing the truth was costfree, so to speak – then it would be good to believe all and only what is true' (Lynch 2009:227)

Kvanvig (2008) makes a similar move. He appeals to the idea of idealization to account for the intrinsic value of true belief. According to Kvanvig, part of the epistemic ideal, or the ideal of rationality (no matter what else it might involve), is believing all truths and avoiding all falsehoods. If so, Kvanvig argues, 'no truth can be pointless enough to play no role at all in the story of what it takes to be cognitively ideal' (2008:209–10).

I do not think this line of response is satisfying for it does not address the question we raised earlier, namely, if we (ordinary epistemic agents) do not value uninteresting or trivial true beliefs, why should we think that they have intrinsic value? The answer that if we were ideal epistemic agents, we would value them is not quite right. We are not and will never be ideal epistemic agents. We will always have limited power, resources and other more urgent concerns than contemplating about what is there. Whatever the epistemic or rational ideal might involve for god-like epistemic agents is irrelevant to what is ideal for us. Even if true beliefs about car numbers have some intrinsic 'value' for an ideal epistemic agent, they are not genuinely valuable, for a human-free 'value' is no value at all.

A more plausible line of defence is to say that, valuing true and only true beliefs even with respect to trivial and uninteresting propositions is valuable for its own sake. Taking up this idea, Lynch (2004b: Chp.8 & 9) argues that true belief has intrinsic value because valuing true belief for its own sake has intrinsic value. He offers several arguments which share the following argumentative structure:

(1) The value of X at least partly depends on the intrinsic value of valuing true belief for its own sake.
(2) X is valuable.
(3) Valuing true belief for its own sake has intrinsic value. (1,2)
(4) If the act of valuing true belief for its own sake has intrinsic value, then true belief is valuable for its own sake.
(5) Therefore, true belief has intrinsic value. (3,4)

Candidates of X, according to Lynch, include *intellectual integrity*, which is a willingness 'to stand for one's own best judgment on any matter of importance – not just when it is convenient to do so' (2004b:133), and *sincerity*, which 'is to be disposed to say what you

believe, with the intention not to mislead' (2004b:153).[8] Suppose intellectual integrity and sincerity are valuable, and their value depends on the intrinsic value of valuing true belief for its own sake, does it follow that true belief has intrinsic value?

I agree with Lynch that valuing true belief for its own sake has intrinsic value, but I do not think it follows that true belief has intrinsic value. This is because *valuing true belief* for its own sake takes more than an act of believing p and refraining from believing not-p when p is true. It involves the manifestation of a pattern of doxastic attitudes, or, again as I would prefer to call it, a *practice* of belief management governed by the truth norm of belief. It does not follow from the fact that valuing true belief for its own sake as a practice has intrinsic value, that every true belief has intrinsic value. What is valuable is the practice of valuing true belief for its own sake, which includes, say, valuing a true belief about car numbers for its own sake, but this is not to say that the *belief* about car numbers is itself valuable in virtue of being true. Therefore, Lynch's appeal to the intrinsic value of intellectual integrity and sincerity does not support the claim that true belief has intrinsic value. Rather, it at best suggests that the practice governed by the truth norm of belief has intrinsic value, which motivates the practice-based account of the truth norm I will defend in the next chapter.

To sum up, we have not found a convincing argument for the claim that true belief has intrinsic value. The argument for the evaluative account by appealing to the intrinsic value of true belief fails to hold water.

2.2.3 True Belief and Basic Desire

Another common approach to support the evaluative account is to appeal to the idea that we have a basic desire for true belief. This remains a popular idea among philosophers ever since Aristotle's famous declaration that men by nature desire to know. Here we have a straightforward argument for the obtaining for the desire-based grounding fact: for any p, if p were true, believing p and refraining from believing not-p would promote desired states of affairs because we have a basic desire for true belief.

But why should we accept the premise that true belief is desired in virtue of its being true? Lynch (2004a) uses what I will call a contrastive strategy to justify the claim that we

---

[8] In a similar vein, Williams (2002) argues that valuing true belief for its own sake constitutes what he calls 'virtues of truth' which include the virtue of 'accuracy' (dispositions to get one's beliefs correctly) and the virtue of 'sincerity' (dispositions to pass on one's beliefs correctly). Both virtues are valuable and instrumental to the flourishing of humankind.

have a basic desire for true belief. He does so by showing that in a wide range of cases, it seems that we prefer true belief over false belief:

> (CASE 1) Suppose we have two beliefs with identical instrumental value, but one is true and the other is false. We prefer the true belief to the false belief. (2004a:502)

> (CASE 2) Suppose we have some false beliefs that are undetected, and it makes no practical difference to one's life whether they are removed or retained. We would prefer to remove those beliefs rather than keep them. (2004a:503)

> (CASE 3) Consider a Cartesian demon world which is experientially indistinguishable from the actual world. None of the true beliefs we have in the actual world are true in the demon world, but they have exactly the same experiential consequences. We would prefer to live in the actual world and have true beliefs about the world rather than live in the demon world and have false beliefs about the world. (2004a:503)

> (CASE 4) Consider a Russellian world that just sprang into existence a second ago which is experientially indistinguishable from the actual world. None of the true beliefs we have in the actual world about the past are true in the Russellian world, but they have exactly the same experiential consequences. We would prefer to live in the actual world and have true beliefs about the past rather than live in the Russellian world and have false beliefs about the past. (2004a:504)

These intuition-pumps seem to suggest that we have a basic preference for true belief, as Lynch (2004a:504) puts it, 'I want my beliefs to track reality, to "accord with how the world actually is" – which is to say I want them to be true.'[9]

[9] It is worth pointing out, in making this point Lynch is importing a correspondence conception of truth.

Suppose we share the intuition that in each case we prefer true belief to false belief as far as the description of the case goes. Nevertheless, I think Lynch's contrastive strategy for establishing a basic desire for true belief fails for two reasons.

First, for the argument to work, Lynch must show that in each case our preference is explained by or indicates a basic desire for true belief. One way to challenge the argument is to point out that there might be alternative explanation of our preference in each case even if a basic desire for truth is lacking. There are numerous reasons why we would prefer to live in the actual world rather than live in the Cartesian demon world or the Russellian world that are not reducible to our desire for true belief. The fact that in CASE 3 and 4, of two experientially indistinguishable worlds we would prefer one to another at best shows that desirable experience is not all we care about if we were to choose worlds. Likewise, there are numerous reasons why we would prefer a true belief to a false belief with the same utility. The fact that in CASE 1 and 2, of two equally useful beliefs, we would prefer to keep one and have another removed at best shows that the utility of belief is not all we care about.

Another problem with Lynch's argument is that our preference for true belief in all four cases is contrastive and conditional. But we cannot simply derive a basic desire for true belief from a preference for true belief to false belief under a certain set of conditions. The move from 'if so and so were the case, we would prefer X to not-X' to 'we prefer X' is invalid. A contrastive preference for X does not entail a desire for X. My preference for vanilla ice cream over strawberry ice cream does not entail a preference for vanilla ice cream *simpliciter*. I dislike both. But if I have to choose, I would rather have the vanilla ice cream than the strawberry ice cream. So, a conditional and contrastive preference for X is compatible with the fact that X is undesired, and it does not entail a desire for X.

Furthermore, the view that true belief is desired for its own sake faces similar difficulty regarding the car numbers case. Imagine a person with a notebook setting out to learn the car numbers because he believes, correctly, there are such truths to be gathered about the car numbers. If, as Lynch claims, there is a basic desire for true belief, then this person's action should be intelligible to us. But we cannot make sense of this person's action.

In response, a parallel move can be made to appeal to the idea of an ideal epistemic agent with unlimited resources and for whom believing is costfree. Perhaps an ideal epistemic agent would want to believe every truth out there, including truths about car numbers.

But this line of response is unsatisfying for the same reason: we are not and will never be ideal epistemic agents. Whatever the epistemic or rational ideal desired by god-like epistemic agents is irrelevant to what is ideal for us.

In short, I do not think Lynch's argument shows that we have a basic desire for true belief. Thus, the argument for the desire-based grounding fact from the idea that we have a basic desire for true belief is unsuccessful.

***

In this section, I have considered three ways to support the claim that the reason to believe p and to refrain from believing not-p when p is true is either value-based or desired-based. But we have seen that none of the arguments is satisfactory. The evaluative account of the truth norm thus fails to meet the grounding challenge.

Let us now turn to a similar concern I raised for the constitutive account. For the constitutive account to successfully meet the grounding challenge, its proponent must show that the relevant constitutive grounding fact obtains, i.e., it is constitutive of belief that if p were true, p would count in favour of believing p and refraining from believing not-p. In the next section, I will consider and reject three arguments for the claim that the constitutive grounding fact obtains.

2.3 Why Truth is not a Constitutive Reason for Belief

2.3.1 Individuating the 'Belief Box'

The first argument for the claim that the constitutive grounding fact obtains is what I will call the argument from individuation:

(1) If some condition C individuates some type of mental state M from non-M, then C is a constitutive condition of M.
(2) That for any p, if p were true, p would count in favour of believing p and refraining from believing not-p is a condition that individuates belief from other mental attitudes.
(3) Therefore, the constitutive grounding fact obtains: for any p, it is constitutive of belief that if p were true, p would count in favour of believing p and refraining from believing not-p.

Assuming that premise (1) follows from our understanding of the constitutive condition, the key premise in this argument is premise (2). The line of thought supporting premise (2) can be traced back to Anscombe (1957:56), who makes the distinction between mind-to-world direction of fit and world-to-mind direction of fit. Belief is the kind of attitude that has a unique 'mind-to-world' direction of fit, that is, what one believes should reflect how the

world is. When your belief does not match the world, the mistake lies with your belief. Intention, by contrast, is a kind of attitude that has a world-to-mind direction of fit. When your intention does not match the world (i.e., it is not realized), the mistake does not lie with your intention.

Williams (1973) makes a similar point. He observes that belief is not merely a propositional attitude that has representational content, for other propositional attitudes such as imagination, hope, fear, desire, etc., also have representational content. He suggests that it is the fact that belief aims at truth that distinguishes belief from other propositional attitudes. When beliefs are true, they are successful. When beliefs are false, they are defective. Truth necessarily counts in favour of belief, but the same does not apply to other mental attitudes.

Both Anscombe and Williams can be interpreted as making the point that it is the normative relation between truth and belief that distinguishes belief from other mental attitudes. One reason why you might find premise (2) attractive is that it fits nicely with the representationalist account of belief we mentioned in Sect. 1.2.

Recall that the representationalist holds that to have a belief p is to stand in a particular relation to an internal mental representation with the content 'p' stored in a 'Belief Box' and poised to perform some specific set of causal roles. The representationalist account of belief is part of the wider representational theory of mind, according to which different mental states can share contents. At the level of thought, for example, the thought that 'water is transparent' and the thought that 'water is tasteless' share a constituent, i.e., 'water'. Constituents like this can be repeated and recombined to produce new thoughts in a systematic and productive way. At the level of attitude, different mental attitudes can share the same content. If you can believe p, then you can also imagine p, fear p, desire p, etc. What individuates belief from imagination, fear, desire, etc., is the particular relation the thinker stands to the internal mental representation of p. To use the well-known metaphor, we may imagine that there is an 'Imagination Box', 'Fear Box', 'Desire Box' alongside the 'Belief Box', where we can place tokens of the same type of mental representation. This picture is so central to the representationalist theory of mind that is sometimes called *boxology*. We need to account for the conditions that individuate the various 'boxes' in our mental economy.

Now, if you are attracted to the representationalist approach, you might also find premise (2) attractive. Premise (2) can be seen as providing a normative account of the 'Belief Box'. Take an attitude you have towards p. What it takes for that attitude to be the 'Belief

Box', is for the following normative condition to obtain: if p were true, p would count in favour of having that attitude towards p and refraining from having that attitude towards not-p.

One objection to premise (2) is that there might be some other mental attitude, such that, for any p, if p were true, p would count in favour of having that attitude towards p and refraining from having that attitude towards not-p. If so, the normative condition is *not sufficient* to distinguish belief from that mental attitude. And thus, it cannot be the condition which individuates the 'Belief Box'. For example, Owens (2003) argues that although the normative condition can distinguish belief from attitudes such as desire and hope, it fails to distinguish belief from guessing. Guesses, just like beliefs, are correct if and only if they are true. Guessing, just like belief, aims at truth. So, we might think that for any p, if p were true, p would also count in favour of guessing p and refraining from guessing not-p.[10]

One might object that guessing is not a distinct propositional attitude from belief. In saying that 'I guess there will be a patent waiver for Covid-19 vaccines', it seems that I am not expressing an attitude distinct from belief. But rather, I am expressing a belief with a relatively low confidence. It may be argued that guessing and belief are not two distinct boxes. Rather, they are both part of the 'Belief Box' but 'held' with different degrees. If so, Owen's argument does not take off.

Our intuition, however, can be pulled in different directions regarding whether guessing is distinct from belief. In paradigmatic cases of guessing such as making a bet, guessing seems to be something quite different from belief. For example, in saying that 'I guess ticket number 9 will be the winner', I seem to express an attitude quite distinct from belief – it is not a belief with a relatively low degree of confidence. I am not at all committed to the truth of the proposition that ticket number 9 will be the winner. I am quite indifferent to the truth of the matter.

I have said that belief comes with high degree of confidence. Guessing, by contrast, need not come with high degree of confidence at all. One might guess p even if one has zero confidence in the truth of p. So, although there might be borderline cases between belief with relatively low degree of confidence and guess with relatively high degree of confidence, belief with high degree of confidence and guess are clearly two distinct attitudes.

---

[10] This thought motivates Owens to make a case for a norm of knowledge as the correct constitutive norm of belief, because he thinks a norm of knowledge, but not a norm of truth can distinguish belief from mere guessing.

Be this as it may, a more pressing objection, I think, is that the normative condition is not well motivated since the condition is not *necessary* to distinguish belief from other mental attitudes. The most common way to individuate the 'Belief Box' is *not* by appealing to a normative condition, but by engaging in functional analysis, i.e., by identifying the relevant causal structure of the type of attitude in a system.

There are, broadly speaking, two approaches. Proponents of *analytic functionalism* engage in such functional analysis by drawing from folk-psychology. For example, a typical analysis of the 'Belief Box' may include the following causal structure, which has perceiving and inferring *p* as inputs (which tends to cause the belief that *p*), and has actions and behavioural dispositions as outputs (which are apt to be caused by possessing the belief that *p* in combination with some desires). Proponents of *psychofunctionalism* engage in functional analysis by drawing on empirical data from cognitive science. On this approach, the correct functional analysis is given in terms of the law-like generalizations uncovered in cognitive science. [11] Either way, the normative condition appears unnecessary – it is the causal/functional roles belief plays in a system that individuate belief from other propositional attitudes.

While the functionalist approach is rooted in empirical sciences and supported by experimental data, premise (2) is a claim grounded in philosophical reflection about the putatively normative nature of belief alone. This is not to say, 'armchair' philosophy has no place in our understanding of the nature of belief. However, it does seem to me that, if we are presented with a plausible functionalist account of the 'Belief Box', supported by empirical evidence uncovered in cognitive science, there is little motivation to prefer the normative account. It is, I believe, up to cognitive scientists to draw up a comprehensive profile of belief that identifies and differentiates belief from other propositional attitudes.

It seems, then, in the absence of further arguments for preferring the normative account over the functionalist one, we must conclude that the normative condition is not necessary for individuating belief from other propositional attitudes. Therefore, the argument from individuation fails to show that the normative condition that for any p, if p were true, p would count in favour of believing p and refraining from believing not-p is constitutive of belief.

---

[11] See, e.g., Quilty-Dunn & Mandelbaum (2018); Porot & Mandelbaum (2021).

2.3.2 Essential Normative Dispositions

In the previous section, the argument from individuation is evaluated in combination with a representationalist/functionalist account of belief. I will now consider a variation of this argument from a dispositionalist perspective, developed by Wedgwood (2007: Chp.7). Wedgwood argues that, if having a belief is just to have a cluster of dispositions characteristic of that type of attitude, then there must be a certain cluster of dispositions that determines the attitude-type as being that of belief. He goes on to argue that the attitude-type of belief is determined by rational dispositions specified by rational principles governing belief which include the truth norm. Let me unpack these claims.

A disposition, as commonly understood, can be characterized as a function mapping input conditions to the corresponding output conditions. For example, solubility can be characterized as a disposition mapping some input condition (e.g., being placed in water) to some output condition (e.g., being dissolved). Anger can be characterized as a disposition mapping some input conditions (e.g., being insulted, being hit) to some output conditions (e.g., face turning red, fists clenching). If an object O has a certain disposition with input conditions Xs and output conditions Ys then the following counterfactual is usually true of O: if O were Xs, then O would be Ys.[12]

According to the *dispositionalist* theory of belief, a belief is a mental attitude that involves a cluster of dispositions standing in relation to other mental states in a complex system. It involves dispositions mapping input conditions such as perceiving or inferring p, to output conditions such as actions or assertions in combination with appropriate desires.[13]

The key to understand Wedgwood's argument is the notion of *rational dispositions*. Rational dispositions, according to Wedgwood, are dispositions whose input conditions essentially involve normative properties and relations specified by relevant rational principles (i.e., norms) governing the attitudes and concepts involved. For example, the disposition to believe that it is raining when one's visual experience constitutes *a reason* for believing that it is raining (e.g., when one perceives that it is raining in the absence of defeaters) counts as a rational disposition.

---

[12] Notice that this is not a counterfactual account of dispositions. The counterfactual is a useful but fallible guide to whether O has the relevant dispositions.

[13] More rigorously, the complex of dispositions can be described by a 'Ramsey sentence' (e.g., Lewis 1970). The idea is that a theoretical term can be eliminated by using an existentially quantified sentence containing observational terms.

Of all these various normative properties and relations specifying rational dispositions involved in belief, Wedgwood thinks the rational disposition specified by the truth norm of belief is the most fundamental one, namely, that p is true is a reason to believe p and to refrain from believing not-p. If belief essentially involves rational disposition to respond to reasons, including the reason to believe p and to refrain from believing not-p when p is true, then the constitutive grounding fact obtains. We can formulate Wedgwood's argument as follows:

(1) Dispositionalism about belief is true: having a belief is just to have a cluster of dispositions characteristic of that type of attitude.
(2) *Rational* dispositions to form and revise an attitude in response to various reasons are characteristic of the attitude-type of belief, of which the most fundamental one is the rational disposition to form that attitude about p and to refrain from forming that attitude about not-p in response to the reason that p is true.
(3) If having a belief is just to have a cluster of dispositions characteristic of that type of attitude, then there must be a certain cluster of dispositions that determines the attitude-type as being that of belief.
(4) What determines the attitude-type of belief must be rational dispositions.
(5) Therefore, belief essentially involves rational dispositions. (1,3,4)
(6) Therefore, the constitutive grounding fact obtains: for any p, it is constitutive of belief that if p were true, p would count in favour of believing p and refraining from believing not-p. (2,5)

Since both dispositionalism about belief and the idea that belief involves rational dispositions are plausible, the key premises are (3) and (4).

Consider premise (3) first. Why should we accept that there must be a precise and specific set of dispositions that determine the attitude-type of belief? Wedgwood argues as follows:

> '… if there is a community that has a type of attitude that they are disposed to form and revise in just the same way as we are disposed to form and revise our beliefs, it would be absurd to suppose that this attitude is in fact the attitude of *disbelief* – so that almost all of their attitudes of this type are incorrect and irrational. Again, something about our dispositions with respect to an attitude-

type must determine exactly which attitude-type it is that figures in our thought.' (Wedgwood 2007:166, his emphasis)

If I understood Wedgwood correctly, the thought seems to be this. There is a clear case in which it is wrong to ascribe an attitude-type other than that of belief to some agent with a certain set of dispositions. To explain this fact, there must be some set of dispositions that determines the attitude-type of belief. For then we can appeal to the presence of that set of dispositions essential to the attitude-type of belief to explain why it is clearly wrong to ascribe an attitude-type other than that of belief in cases such as the one Wedgwood invites us to consider.

But there is an alternative, simpler and more plausible explanation why it is wrong to ascribe the attitude-type of disbelief to agents of that community, namely, they manifest dispositions *characteristic* of the attitude-type of belief and they do not manifest dispositions *characteristic* of the attitude-type of disbelief. The appeal to the characteristic dispositions, i.e., dispositions that are usually involved in belief is sufficient to explain the above case. So, Wedgwood's argument fails to support his claim that there *must* be a precise and specific set of dispositions that determine the attitude-type of belief.

Moreover, this claim conflicts with one of the central motivations of the dispositionalist account of belief in the very first place. By identifying belief with a loose cluster of dispositions characteristic of belief (i.e., without committing to any specific set of dispositions as essential), dispositionalism has the flexibility to accommodate many borderline cases of belief and explain why a variety of belief-like attitudes can be counted as belief.[14] So by committing to the claim that there must be a precise and specific set of dispositions that determine the attitude-type of belief, Wedgwood's account of belief loses the explanatory strength dispositionalism would otherwise enjoy over alternative accounts of belief.

But even if we accept the idea that some dispositions characteristic of the attitude-type of belief must be essential, why should we think, as claimed in premise (4), that it must be

[14] For example, Schwitzgebel (2001, 2002, 2010, 2013) presents cases of what he calls 'in-between believing' and argues that dispositionalism, but not representationalism can account for such cases. Typical in-between cases involve a subject manifesting some characteristic dispositions we associated with belief but not others.

the cluster of rational dispositions which determines the attitude-type of belief as opposed to some other cluster of dispositions? Wedgwood argues as follows:

> '…any concepts that you have could be shared by a perfectly rational being who had no irrational dispositions at all…so that perfectly rational being would have to possess those concepts in virtue of some of her rational dispositions (since she has no irrational dispositions). This makes it plausible that we also possess those concepts in virtue of some of our rational dispositions, and not in virtue of our irrational dispositions…If this is so, then any adequate account of a concept (or of an attitude-type) must mention this specific rational dispositions, which is essential to possessing that concept of being capable of that attitude-type.' (Wedgwood 2007:168-9)

The thought seems to be this: since we can conceive of a perfectly rational being who is capable of having beliefs, and since that perfectly rational being would not have any irrational disposition (she is disposed to respond to reasons only), any belief she has can only involve a cluster of *rational* dispositions. So, Wedgwood concludes that we are capable of having the attitude-type of belief in virtue of sharing some of our rational dispositions with the perfectly rational being.

Furthermore, Wedgwood contends that *irrational* disposition cannot be involved in determining one's possession of a concept or attitude-type because 'the possession of a concept [or attitude-type] is a cognitive *power* or *ability* – not a cognitive defect or liability' (2007:169). If the possession of a concept or attitude-type is partly determined by some irrational disposition, then irrationality must be involved in every use of that concept and in having that attitude-type. But that seems wrong. On Wedgwood's view, belief, a cognitive power, cannot be determined by irrational dispositions, so it must essentially involve only rational dispositions.

If my interpretation is broadly along the right lines, Wedgwood's argument for premise (4) is unconvincing. First, the argument over-generalizes and leads to unacceptable consequences. Wedgwood attempts to show that the possession of belief must be in virtue of some rational dispositions because a perfectly rational being has the capacity to believe while lacking any irrational dispositions. By the same reasoning, however, any capacity we share with a perfectly rational being would also have to be explained in virtue of some rational dispositions. Take the example of playing a musical instrument. Presumably, my capacity to play the piano involves a cluster of dispositions to move my fingers in certain ways on the

piano in response to the sound I hear, the touch of the keyboard I feel and so on. But there can be a perfectly rational being who also plays the piano. Does it follow that, by the same reasoning, that I possess the ability to play the piano in virtue of some rational dispositions? But it seems too demanding if not outright absurd that my ability to play the piano essentially involves rational dispositions.

Perhaps, Wedgwood could insist that his argument only applies to the capacity of concepts possession and attitude-types, and therefore the piano case misses the point. But this line of response seems *ad hoc*. Why is it the case that some of our capacities shared by a perfectly rational being have to be determined by rational dispositions but some other capacities shared by a perfectly rational being need not? Why is it the case that concept-possession and attitude-types belong to the former while capacities such as playing an instrumental belong to the latter?

Furthermore, the underlying assumption of Wedgwood's argument is that every disposition a perfectly rational being has must be rational. But, plausibly, many dispositions are neither rational nor irrational. For instance, dispositions involved in playing a musical instrument do not appear to be either rational or irrational. If so, Wedgwood cannot, without begging the question, assume that the possession of belief for a perfectly rational being must be in virtue of some of her rational dispositions, rather than in virtue of some dispositions that are neither irrational nor rational. It may very well be that believing p, like playing the piano, is determined by dispositions that are neither rational nor irrational. For this reason, even if Wedgwood is right that irrational dispositions cannot be involved in determining the attitude-type of belief (because it is a cognitive power), it does not follow that it must be some rational dispositions in virtue of which an attitude is that of belief.

Having said this, there may be a separate way to support premise (4). Nolfi (2015) develops an argument which may be employed to support premise (4). She invites us to consider the following case:

> '…take a creature that clearly has the capacity for belief. Strip this creature of all and only dispositions to regulate and employ her mental attitudes in ways that would either constitute or approximate rational belief formation and employment.' (Nolfi 2015:184)

According to Nolfi, we would not regard this creature as a believer. The stripping of rational dispositions characteristic of belief is so radical that post-transformation, 'it makes no sense to say that this creature's mental attitudes are of the same kind as the mental attitudes of

paradigmatic believers' (2015:184). Therefore, the rational dispositions we share with paradigmatic believers are *essential* to what it is to have the attitude-type of belief.

As I have said earlier, I agree with Wedgwood that rational dispositions are characteristic of the attitude-type of belief. It is characteristic of having a belief p that one usually has the relevant rational dispositions, for instance, to believe p and to refrain from believing not-p in response to the reason that p is true. But is it true that if we strip a creature of all and only rational dispositions, her mental economy would be radically different from that of paradigmatic believers such that it is unintelligible to ascribe the attitude-type of belief to that creature? Crucially, for our purposes, is it the case that if someone lacks the disposition to believe p and to refrain from believing not-p in response to the reason that p is true, that person would have a radically different mental economy from that of paradigmatic believers?

I think not. We can imagine a person gaining a disposition to believe p and to refrain from believing not-p in response to the fact that believing p is useful, but is deprived of all rational dispositions involved in belief including the disposition to believe p and to refrain from believing not-p in response to the reason that p is true. As it happens, useful beliefs are often true. We would expect that, post-transformation, this person would have roughly the same belief-stock and exhibit similar behavioural pattern as before. Thus, contrary to what Nolfi claims, post-transformation, it may still be intelligible to ascribe the attitude-type of belief to that person since her mental economy would not radically diverge from that of paradigmatic believers.

To block this type of response, Nolfi could say that the disposition to respond to the utility of belief, though not itself rational, approximates rational dispositions involved in belief and should be stripped as well. But by making this move, the argument loses its force. For in that case, the radical divergence of the mental economy of the person post-transformation may be attributed to the fact that she does not have dispositions that *approximate* rational dispositions but are not themselves rational dispositions. So, it does not follow that our possession of belief must be in virtue of some of our rational dispositions shared with paradigmatic believers.

Alternatively, Nolfi could press on and insist that the person who has a disposition to respond to the utility of belief but is deprived of all rational dispositions has a radically different mental economy from paradigmatic believers. After all, post-transformation, this person would tend to believe what is convenient and pleasant to believe. But this does not seem to be a radical departure from the mental economy of paradigmatic believers. There is ample empirical evidence suggesting that the way we are disposed to form beliefs is often

insensitive and irresponsive to evidence but motivated by our desires and wishes. Irrational belief-formation is quite pervasive in our daily life. The psychological mechanisms underlying irrational belief-formation such as cognitive dissonance, implicit bias, etc., are well-documented.[15] We will return to this point and have a closer look of the empirical literature on belief-formation in Chapter 4. But for now, it suffices to say that Nolfi's argument for premise (4) by appealing to paradigmatic believers is unpersuasive.

To sum up, since we have no good reason to accept either premise (3) or (4), it seems, we can conclude that Wedgwood's argument from individuation fails to establish the constitutive grounding fact with or without Nolfi's supporting considerations. Although rational dispositions are characteristic of belief, there is no sufficient support for the claim that they are essential. In fact, as I will argue later in Chapter 4 & 5, one's belief is *guided* by epistemic norms if one manifests such rational dispositions to respond to reasons specified by the norms in question. Since our beliefs are not always guided, naturally, not all our beliefs involve rational dispositions. Rational dispositions are not essential to unguided beliefs.

2.3.3 Transparency and the Concept of Belief

One lesson we can extract from the failure of both variations of the argument from individuation is that the constitutive grounding fact cannot be established as a *metaphysical* fact about belief being the type of attitude-type it is. This leads us to an alternative interpretation of the constitutive grounding fact as a claim about the *concept* as opposed to the *ontology* of belief. To grasp the concept of belief, is to grasp the fact that for any p, if p were true, p would count in favour of believing p and refraining from believing not-p. According to this interpretation, it is constitutive of the *concept* of belief that if p were true, p would count in favour of believing p and refraining from believing not-p.

One influential argument for this conceptual claim is what I will call the argument from doxastic deliberation (Shah 2003; Shah & Velleman 2005). According to Shah and Velleman, doxastic deliberation is an act one engages in when one deliberates about whether to *believe* a proposition p. Doxastic deliberation has a feature that they call 'transparency', that is, the deliberative question whether to believe p gives way to the factual question whether p is true, to the exclusion of any other questions; and the thinker takes the considerations that answer the factual question whether p is true to be the considerations that answer the deliberative

[15] For an overview of recent studies on irrational belief-formation, see, e.g., Porot & Mandelbaum (2021).

question whether to believe p, to the exclusion of any other considerations (Shah 2003:447, 463; Shah and Velleman 2005:499, 501). They go on to argue that transparency is best explained by the fact that it is constitutive of the concept of belief that belief is subject to a norm of truth. When one reasons about whether to believe p, one deliberates about an attitude, i.e., belief. One therefore necessarily employs the concept belief to which one applies the truth norm of belief and grasps the fact that it is the truth of p that counts in favour of believing p. Therefore, the question whether to believe p gives way to the question whether p is true and the deliberating agent takes the considerations that answer the factual question whether p is true to be the only relevant considerations that answer the question whether to believe p. We can reformulate their argument in the present context as follows:

(1) Doxastic deliberation is transparent.
(2) Transparency is best explained by the constitutive grounding fact: it is constitutive of the *concept* of belief that if p were true, p would count in favour of believing p and refraining from believing not-p.
(3) Therefore, the constitutive grounding fact obtains.

Shah and Velleman take premise (1) to be a self-evident truth and the starting point of their argument. On one natural interpretation of what Shah and Velleman might mean when they say that doxastic deliberation is transparent, transparency is part of the phenomenology of doxastic deliberation. It is a psychological fact about how we deliberate whenever we ask ourselves whether to believe p. For anyone who is deliberating about whether to believe p, that person would experience a 'seamless', 'immediate' and 'non-inferential' move to the question whether p is true and take considerations answering the latter question to be the only relevant considerations that settle the former question (e.g., Shah 2003:447).

However, it seems to me there are clear counterexamples to transparency. When asking myself whether to believe that God exists, that question did not immediately give way to the question of whether God exists is true. Rather, it was settled by my belief that the question is meaningless. When deliberating about whether to believe that my funding application will be successful, that question did not give way to the question of whether it is true that my funding application will be successful. Rather, it was settled by my belief that it does me no good to speculate about the result. When deliberating about whether to believe that my friend will become a successful musician, that question did not give way to the factual question of whether it is true that my friend will become a successful musician. Rather, it was settled by

my belief that I shall have faith in my friend regardless of the odds. The list of examples goes on.

Furthermore, it seems wrong to say that the two questions 'whether to believe p' and 'whether p is true' are unseparated. This is simply not the way we reason and for good reasons. As critical thinkers, in deliberating about whether to believe p, we must not have our eyes exclusively on the truth of p and leave aside other important and relevant questions. For instance, does the articulation of the proposition p involve problematic concepts? Am I in a suitable epistemic position to ask myself whether to believe p? Is now the best time to deliberate about whether to believe p? What is there to care about whether to believe p or not?

In short, it is easy to conceive of someone who is deliberating about whether to believe p but during the process of her deliberation of the question neither gives way to the factual question whether p is true, nor is settled by considerations that bear exclusively upon the truth of p. Transparency, at least in its present form, is clearly a false description of how we actually deliberate. Since doxastic deliberation is not transparent, there is little motivation to further engage with premise (2) here.[16]

But perhaps, there is a different interpretation of what Shah and Velleman might mean when they say that doxastic deliberation is transparent. On this interpretation, transparency is not a feature about how we deliberate, empirically speaking, when we ask ourselves whether to believe p, but a claim about what it is to engage in doxastic deliberation proper. If doxastic deliberation is *necessarily* transparent, then in the above cases, either one is not deliberating about an attitude of belief or one does not grasp the concept of belief properly.

But this raises the question about what one wants to achieve with the argument from doxastic deliberation. Critics of the constitutive account of the truth norm, who challenges its proponents to say more about why the constitutive grounding fact obtains, are unlikely to accept the premise that doxastic deliberation is necessarily transparent. They will not be impressed by an argument which says that the transparency of doxastic deliberation *proper* is best explained by the constitutive grounding fact. They will not accept, without further arguments, the claim that doxastic deliberation is necessarily transparent.

---

[16] See McHugh (2013) for further critical discussion of premise (2). Roughly, McHugh argues that, Shah and Velleman's argument for premise (2) relies on a strong version of motivational internalism, which says that the acceptance of the truth norm of belief necessarily motivates the thinker to reason in accordance with that norm, in this case, to be moved solely by evidential considerations. But it is dubious that strong motivational internalism is true.

To sum up. The argument from doxastic deliberation does not go through on either interpretation of what Shah and Velleman might mean when they say that doxastic deliberation is transparent. Therefore, the proponent of the constitutive account is yet to convince us that the constitutive grounding fact obtains. In the absence of such arguments, the constitutive account fails to meet the grounding challenge.

2.4 Conclusion

In this chapter, I first provided a brief overview of the two main views in the current literature about the authority of the truth norm: the evaluative account and the constitutive account. I outlined some existing problems for both views. I then identified and explained what I take to be a more fundamental concern for both views, which has not received sufficient discussion. On the reason-based normative framework, the fundamental challenge to both the evaluative and the constitutive account is to show that the relevant reason grounding facts obtain. I then considered three arguments for the claim that the reason to believe p and to refrain from believing not-p when p is true is value-based/desired-based and three arguments for the claim that the reason is constitutive. But none of the arguments is satisfactory. Thus, proponents of the evaluative and constitutive account of the truth norm fail to meet the grounding challenge since they are yet to convince us that their preferred reason grounding facts obtain.

3. Defending a Practice-based Account of the Truth Norm

In the previous chapter, I argued that both the evaluative and the constitutive account of the truth norm fail to meet the grounding challenge. The grounding challenge, recall, concerns whether the reason to believe truly and only truly can be properly grounded. My aim in this chapter is to develop and defend a practice-based account of the truth norm. According to this view, the authority of the truth norm of belief is grounded in what I will call the *T-practice*, a justified social practice governed by the truth norm of belief. The reason to believe truly and only truly is practice-based.

To develop and defend the practice-based account of the truth norm, I will first motivate the view by contrasting it favourably with the evaluative and the constitutive account (Sect. 3.1). For the practice-based account of the truth norm to successfully meet the grounding challenge, I must show that the practice-based grounding fact obtains. To this end, I will first suggest that a social practice has three earmarks (Sect. 3.2). I will then argue that the T-practice is a social practice satisfying all three earmarks (Sect. 3.3). Next, I will contend that the T-practice functions to facilitate knowledge production, maintenance, and social cooperation. I will argue that the T-practice is a justified social practice since its social function is indispensable to the survival and flourishing of human societies, and even a partial breakdown of the T-practice would cause substantial harm. The reason to believe p and to refrain from believing not-p when p is true is practice-based (Sect. 3.4).

3.1 Motivating the Practice-based Account of the Truth Norm

3.1.1 The Social Dimension of Belief

Despite the rapid development of social epistemology in recent decades, both the normativists and the anti-normativists about the truth norm of belief have not paid much attention to the social dimension of this debate.[1] One important motivation for the practice-based account of the truth norm is that it takes account of the social aspect of belief: that beliefs are transmitted and shared among epistemic agents living in epistemic communities.

[1] In contrast to the traditional Cartesian epistemology, which focuses on how an individual comes to have knowledge through perception and reasoning, social epistemologists have come to acknowledged the complex social web in which knowledge (and belief) acquisition and transmission take place (e.g., Pritchard et al. 2008; Goldman & Whitcomb 2010; Fricker et al. 2019).

To illustrate the distinctive sense in which beliefs are social, consider a few examples. I believe a lot of things for which I do not myself possess evidence: that Jupiter has 79 moons, that the Pfizer/BioNTech vaccine is effective against Covid-19, that killer whales produce a variety of sounds for navigation and communication, that the battle of Austerlitz took place in 1805, that mashing is the first step in the beer brewing process, that Fermat's last theorem has been proved… I acquire a large, if not (potentially) infinite, number of beliefs by learning from other epistemic agents. For some of the beliefs I acquire from others, I could go and check the evidence supporting the truth of those claims by myself – I could perhaps easily find out for myself what the first step for beer-brewing is. But for many beliefs, I do not have direct access to the relevant evidence and even if I come into possession of such evidence, I may not be able to evaluate them properly. For instance, I rely on the community of historians for obtaining evidence about the battle of Austerlitz. I rely on the community of mathematicians for their evaluation of the validity of Wiles's proof of Fermat's last theorem, for I lack the ability to do so myself.

We live in an epistemic community, a social network made up of epistemic agents like you and me who are capable of acquiring information and forming beliefs. Each epistemic agent makes use of the beliefs of others to manage her own beliefs through various forms of interactions. Belief has an important social dimension in the sense that it is something that can be shared and transmitted in social networks.

Now, for beliefs I acquire from other epistemic agents, I rely on the producers of those beliefs to adhere to certain standards of production and maintenance. For beliefs I acquire for myself, I am also expected to adhere to certain standards of production and maintenance so that they can be shared with other epistemic agents in my epistemic community. One of the standards, I submit, is the truth norm of belief. We have a responsibility towards one another that we manage our beliefs correctly.

The Cartesian paradigm of a self-sufficient and self-reliant epistemic agent forming and revising her beliefs in isolation is maybe the right focus of evaluation in some epistemic inquiries, but it is, I believe, not the right starting point to theorize about why the truth norm of belief is authoritative. This perspective motivates us to treat our belief management not as an isolated practice that each individual engages in on her own, but as a social practice that we engage in together as a community. If this is right – that belief management can be reasonably characterized as a social practice, then we have some initial motivations for thinking that the reason to believe p and to refrain from believing not-p when p is true could be practice-based.

3.1.2 The Wrong-kind-of-reason Problem

Another motivation for the practice-based account of the truth norm is that it avoids the wrong-kind-of-reason problem facing the evaluative account of the truth norm. Recall that, the evaluative account seeks to meet the grounding challenge by appealing to this intuitive idea that true belief is desirable/valuable. But the evaluative account of the truth norm seems to offer the wrong kind of explanation of why truth is a reason for belief. For it seems that the truth of p is a reason to believe p and to refrain from believing not-p even if that belief is not desirable/valuable. Furthermore, it seems to force one to commit to the idea that there are other desire-based/value-based reasons for belief. There are many ways in which a belief might promote desired/valuable states of affairs regardless of whether it is true or not. But many philosophers take desire-based/value-based reasons to be the wrong kind of reasons for belief.

The practice-based account, by contrast, avoids the wrong-kind-of-reason problem. It seeks to meet the grounding challenge by appealing to the idea that believing truly and only truly constitutes a justified social practice, i.e., the T-practice. If believing p and refraining from believing not-p when p is true would constitute a justified social practice, then the practice-based grounding fact would explain why the truth of p is a reason to believe p and to refrain from believing not-p. The T-practice-based reason to believe p and to refrain from believing not-p is distinctive of the attitude type of belief. It does not force one to accept other desire-based or value-based reasons for belief and thus avoids the wrong-kind-of-reason problem.

3.1.3 The Problem of Limited 'Authority'

The practice-based account of the truth norm also avoids the problem of limited 'authority' facing the constitutive account of the truth norm. Recall that, the constitutive account seeks to meet the grounding challenge by appealing to this intuitive idea that truth is the constitutive norm of belief. It is constitutive of belief that the truth of p counts in favour of believing p and refraining from believing not-p. So, as far as you are playing the game of belief, the truth of p is a reason for you to believe p and to refrain from believing not-p. But the constitutive account faces the problem of limited 'authority'. It seems that, for anyone who is not playing the game of belief with respect to p, one has no reason to believe p and to refrain from believing not-p when p is true. But this seems wrong, for it seems that the

truth of p is a reason to believe p and to refrain from believing not-p regardless of whether one is playing the game of belief or not.

The practice-based account, by contrast, avoids the problem of limited 'authority'. What grounds the reason to believe p and to refrain from believing not-p when p is true is the fact that the T-practice is a justified social practice, regardless of whether one is playing the game of belief. But you might wonder, isn't there a similar worry for the practice-based account? Isn't it true that, for those who do not play the game of the T-practice, the truth of p is not a reason to believe p or to refrain from believing not-p?

This does not need to worry us. Even for individuals who are not playing the game of the T-practice, the practice-based reason is nonetheless normatively binding. This is because as far as the T-practice is a justified social practice operating in their community, the truth of p is a practice-based reason to believe p and to refrain from believing not-p. And as we shall see in Sect. 3.3, the T-practice is a justified social practice operating in our communities because the T-practice is indispensable for the survival and flourishing of human societies. As such, no individual epistemic agent lives outside the 'authority' of the T-practice. So, the practice-based account avoids the problem of limited 'authority'. The truth of p is a reason to believe p and to refrain from believing not-p, even if one is neither playing the game of belief nor playing the game of the T-practice – the truth norm of belief remains authoritative and binding.

***

The above considerations, I believe, offer good motivations for the practice-based account of the truth norm. But recall that, in the previous chapter I identified a concern for the evaluative and the constitutive account of the truth norm, which I take to be fundamental on the reason-based normative framework. For their views to successfully meet the grounding challenge, proponents of the evaluative and the constitutive account must show that the relevant reason grounding facts obtain. We considered several ways to establish the desire-based, value-based, and constitutive grounding facts but found none of them satisfying. That, I argued, is why both views fail to meet the grounding challenge.

The same concern arises for the practice-based account. For the practice-based account to successfully meet the grounding challenge, I must show that the practice-based grounding fact obtains. This is the task I will take on in the remainder of this chapter. To show that the practice of believing p and refraining from believing not-p when p is true (i.e., the T-practice) is a justified social practice, I will begin by introducing the notion of social practices.

3.2 Three Earmarks of Social Practices

Practices, as commonly understood, refer to regular patterns of behaviour and attitude exhibited by their participants. Playing tennis, drinking coffee after each meal, working five days a week, celebrating the Spring Festival are in this sense practices. But not all practices count as social practices. Drinking coffee after each meal is usually not treated as a social practice while playing tennis, working five days a week, and celebrating the Spring Festival are. This raises the question of what makes a practice a social practice and whether the T-practice can be reasonably characterized, by that understanding, as a social practice. In this section, I will suggest that a social practice has three earmarks. Our discussion will set the groundwork for my defence of the practice-based account of the truth norm in the next two sections.

3.2.1 The First Earmark: Social Practices Are Rule-governed

Consider the differences between the example we intuitively take to be an individual practice and those we take to be social practices. The first thing to notice is that to engage in a social practice is not simply for participants of that practice to exhibit shared, regular patterns of behaviour. The fact that you, me, and some other likeminded coffee enthusiasts regularly drink coffee after each meal does not make it a social practice. Drinking coffee after each meal, without further qualification, is something an individual does for reasons of her own and is deprived of shared social meaning.[2] So, what makes some shared patterns of behaviour a social practice?

One popular answer among philosophers is to appeal to a rule-governed conception of social practices (see, e.g., Rawls 1955; Brandom 1994; Haslanger 2018).[3] Playing tennis, working five days a week and celebrating the Spring Festival, are things participants engage in together in some communities while adhering to some mutually acknowledged rules governing those practices. Importantly, rules governing a practice need not always be explicit or strictly defined like the rules of tennis. Some rules may be implicit in ways participants of that practice engage one another. For instance, practices of festivities tend to be governed by rules that are more loosely defined, more open to interpretation compared to the rules of

[2] I am not ruling out that in some cultures or communities the practice of drinking coffee after each meal may be counted as a social practice.

[3] Let me be clear here: rules need not be norms for not all social practices are justified. The practice-based account of the truth norm does not beg the question against the anti-normativist by appealing to a rule-governed conception of social practices.

tennis. Implicit rules can evolve over time to reflect various changes in the patterns of behaviour exhibited by participants of that practice.

The first earmark of social practices, I suggest, is that social practices are rule-governed. Participants of a social practice exhibit shared and regular patterns of behaviour in accordance with rules mutually acknowledged by participants of that practice, which may be implicit or explicit.

3.2.2 The Second Earmark: Social Practices Are Interactive

The rule-governed conception of a social practice naturally leads us to another earmark of social practices. If participants engage in a rule-governed social practice, then we would expect them to hold one another mutually responsible to act in ways conforming to the relevant rules. Take a practice of festivities. Participants not only are expected to exhibit patterns of behaviour conforming to the implicit rules governing that practice, but also to engage, interact and respond to one another's behaviour. They are expected to approve and reward behaviours that are in accordance with the implicit rules and to disprove and sanction behaviours that are in violation. By contrast, coffee enthusiasts who engage in the practice of drinking coffee after each meal, without further qualification, are not expected to exhibit such patterns of interaction.

One distinctive type of interactive behaviour exhibited by participants of a social practice is the act of giving and asking for (normative) reasons when participants perceive irregularities, deviations or violations in that practice.[4] Participants of a social practice attribute obligations and permissions to themselves and other participants of the practice in accordance with their understanding of the relevant rules governing that practice. But participants do not always meet one another's expectations about what they are required or permitted to do in a given situation. It may be due to a disagreement about what is the correct

---

[4] My characterization of the second earmark of social practice draws on what Brandom's (1994) famously labels as the *game of deontic scorekeeping*. According to Brandom, to engage in a social practice is not simply for participants of that practice to exhibit shared, regular patterns of behaviour. Crucially, they must engage in a game of deontic scorekeeping. To play the game of deontic scorekeeping, is for participants of that social practice to keep track of their own and other participants' commitments (things one is obliged to do) and entitlements (things one is permitted to do) in accordance with the rules governing that practice. Participants attribute or withhold commitments and entitlements towards themselves and one another, taking some performances to be appropriate or inappropriate. They also engage one another, adjusting their attributions of commitments and entitlements considering what others do and say and be responsive to the relevant demands and expectations from others. The implicit scorekeeping can be made explicit by giving and asking for reasons.

application of the relevant rules in a given situation. Or perhaps they agree about how to apply the relevant rules in a given situation, but some nevertheless fail to do the right things. In such cases, participants of a social practice are expected to engage in the act of giving and asking for reasons to negotiate their differences and bring one another into mutual conformance. The participant who fails to meet other participants' expectations would either be committed to alter her behaviour to meet their demand or defend her behaviour with reasons.

The act of giving and asking for reasons is often accompanied by what Strawson (1962) famously called 'reactive attitudes'. Reactive attitudes, as I understand them here, are attitudes of approval and disapproval participants hold towards themselves and other participants of that practice. Participants tend to exhibit attitudes of approval towards someone (including themselves) when they perceive that agent as conforming to the relevant rules governing the practice, which may include praise, satisfaction, gratitude, and other affirmative attitudes. Participants tend to exhibit attitudes of disapproval towards someone (including themselves) when they perceive that person as violating the relevant rules governing the practice, which may include blame, guilt, remorse, resentment, disappointment, anger, and other critical attitudes.

The second earmark of social practices, I suggest, is that social practices are interactive. Participants of a social practice exhibit a distinctive cluster of behaviour and attitude designed to hold one another mutually responsible for conforming to the rules governing that practice.

3.2.3 The Third Earmark: Social Practices Have Social Functions

What underpins the interactive conception is that a social practice typically has a social function which has a practical impact on its participants. The example we intuitively take to be an individual practice, e.g., the practice of drinking coffee after each meal, without further qualification, does not seem to have any social function, although it may serve some individual needs. By contrast, examples we intuitively take to be social practice such as playing tennis, working five days a week, and celebrating the Spring Festival, function to serve some collective social ends. Crudely speaking, for instance, we might say that the function of the practice of playing tennis is to organize the way we spend our leisure, to produce and distribute pleasure and good health.[5]

---

[5] This, of course, is speculative and my main point here is that social practices typically function to satisfy some collective ends of participants who engage in those practices.

But of course, it is not always possible to read off the function of a social practice directly from a foray into the effects of a paradigmatic instance of that practice. Given the complexity of many social practices, it is not easy to identify the exact function they serve. In Sect 3.4.1, I will introduce a genealogical method to uncover the function of a social practice. But for now, let me just say that a third earmark of social practices is that a social practice typically produces, distributes, or organizes some social goods which have a practical impact on the participants of that practice.

Importantly, not all social practices can be justified by their social functions and many of them should be reformed or abolished altogether. This raises the difficult question of what it is for a social practice to count as a justified practice. I do not have a full answer here. But reflecting upon some examples might help. History is full of examples of unjustified social practices. Some social practices discriminate and oppress particular groups of people, e.g., the practice of slavery, patriarchal gender roles, voter suppression, etc. Some social practices exploit workers and natural resources, e.g., the practice of child labour, unregulated farming, deforestation, etc. An unjustified social practice performs its social function at the cost of causing substantial harm to participants involved in that practice.[6] Unjustified social practices have no authority to ground any normative reasons. There can be no practice-based reason to φ when φ-ing does not constitute a justified social practice. In Sect. 3.4, I will argue that the T-practice is a justified social practice for it functions to promote knowledge production, maintenance and social cooperation which is indispensable to the survival and flourishing of human societies while even a partial breakdown of the T-practice would cause substantial harm.

The third earmark of social practices, I suggest, is that social practices have social functions such as producing, distributing, or organizing social goods which have a practical impact on the participants of those practices. But not all social practices are justified. Only practices which perform their social functions without causing substantial harm to participants involved in those practices can count as justified social practices.

***

We can summarize the points I made as follows: social practices are rule-governed patterns of interactive behaviour and attitude exhibited by a group of agents for producing, distributing, or organizing some social goods. With this characterization of social practices

[6] Causing substantial harm to the environment and animals may also make a social practice unjustified. But for our purposes, I will just focus on the practical impact a social practice has on people.

in place, in the next section I will argue that the T-practice is a social practice satisfying all three earmarks.

3.3 Why the T-practice is a Social Practice

Is there any good reason to characterize the T-practice, i.e., the practice of believing p and refraining from believing not-p when p is true, as a social practice? Let us go through the three earmarks in turn.

First, the T-practice is rule-governed. It is part of our understanding of the T-practice that it involves patterns of belief-formation and revision governed by the truth norm of belief. The practice of believing p and refraining from believing not-p when p is true is a shared pattern of mental behaviour in accordance with the truth norm of belief, which is mutually acknowledged by ordinary epistemic agents like you and me.[7]

Second, the T-practice is interactive. It is common for epistemic agents to exhibit interactive behaviour and attitude toward one another to hold one another mutually responsible for conforming to the truth norm of belief. We attribute epistemic obligations and permissions to ourselves and other epistemic agents in accordance with our understanding of the correct application of the truth norm of belief in a given situation.[8] But, of course, we do not always meet one another's expectations. We often fail to believe the correct things. Sometimes we do not meet one another's expectations because we may disagree about whether a proposition is true, or whether we have an obligation or merely a permission to believe that proposition. In such cases, we engage in the act of giving and asking for reasons to negotiate our differences with the aim of bringing one another into conformity with the truth norm of belief. The epistemic agent who fails to meet others' expectations would either be committed to alter her belief to meet their demands or defend her belief with reasons. It is also common for epistemic agents to respond to one another with various reactive attitudes depending on their judgments about how well the others' belief management conforms with the truth norm of belief.

---

[7] The fact that the truth norm of belief is widely accepted as a self-evident platitude is a good indicator. As I have said, the truth norm needs not be explicitly endorsed for one to be counted as engaging in the T-practice. This allows unreflective epistemic agents to be counted as participants of the T-practice.

[8] This is not to say that we always agree about the correct applications are in a given situation. In fact, philosophers have quite different views about the sort of claim the truth norm can make on us regarding what to believe. I will return to this issue in Part III.

Third, the T-practice has a social function. In our previous discussion of the evaluative account of the truth norm, we have encountered a couple of ideas which provide initial support the claim. For example, Lynch's argument for the claim that true belief is valuable (by appealing to the values of integrity and sincerity) may be seen as revealing the value of the T-practice (Sect. 2.2.2). However, as I have noted, more work is required to uncover the function of the T-practice and I will turn to this task in the next section. For now, it suffices to note that our previous discussion provides some initial reason for thinking that the T-practice satisfies the third earmark of social practices.

However, you might wonder, isn't it true that we have many false beliefs and lack beliefs about many truths? If, in many cases, when p is true, one does not believe p or refrain from believing not-p, then in what sense is the T-practice a shared *regular pattern* of behaviour exhibited by ordinary epistemic agents?

There are two things to say in response. First, we should distinguish the question of whether the T-practice can be reasonably characterized as a social practice from the question of whether the T-practice is an actual social practice most communities engage in. The question concerns us here is the former not the latter. To justify the claim that the T-practice is a social practice, it suffices to show that, the T-practice can be reasonably characterized as a practice involving rule-governed patterns of interactive behaviour and attitude exhibited by epistemic agents for some collective ends.

This, of course, is not to say that the latter question is not important. I do think the T-practice is a common epistemic practice most epistemic agents engage in.[9] This leads us to my second point. It is important to bear in mind that the exhibition of irregularities, deviation or violation of rules are to be expected in a social practice. As I said in Sect 3.2.2, irregularities, etc., are expected to happen and trigger interactive patterns of behaviour (e.g., giving and asking for reasons) and attitude (e.g., reactive attitudes such as blame), which I identified as an earmark of social practices. As such, irregularities, etc., are in fact evidence of an underlying social practice.

The T-practice is not an exception. The fact that we have many false beliefs and lack beliefs about many truths does not imply that the T-practice is not an actual practice ordinary epistemic agents participate in. Overall, the way we manage our beliefs and interact with one another overwhelmingly supports the claim that the T-practice is a common epistemic

[9] In fact, as I will argue in Sect. 3.4, the survival and flourishing of human societies hinge on the T-practice. The T-practice is not something we can be even partially dispensed with without paying a hefty price.

practice we engage in. Note here that engaging in a social practice does not mean that we must be good at it. People break rules in games of football (knowingly and maybe unknowingly) all the time, some are better than others, still there is an underlying social practice.

Having said that, if the irregularities become widespread and persistent, and if a large number of members of a community continuously fail to exhibit rule-governed patterns of interactive behaviour and attitude, then we should perhaps refrain from attributing the T-practice to the community in question. Indeed, I think a partial breakdown of the T-practice can occur in some communities and cause substantial harm, a point I will discuss in more detail shortly in Sect. 3.4.3. How exactly we should draw the line between a deviation and a breakdown of a social practice is a tricky question. But fortunately, it needs not concern us here. To meet the grounding challenge, our goal is to show that the T-practice is a justified social practice and can properly ground the reason to believe p and to refrain from believing not-p when p is true.

To sum up, in this section I argued that the T-practice satisfies all three earmarks of social practices: it is rule-governed, interactive and has a social function. The T-practice is thus a social practice. I will now turn to the question concerning the justificatory status of the T-practice.

3.4 Why Truth is a Practice-based Reason for Belief

3.4.1 The Method of Genealogy

Before we can evaluate whether the T-practice can be justified by its social function, we must first say in more detail what the function of the T-practice is. I will begin by introducing the method of genealogy: a method to uncover the function of a social practice.

To uncover the function of a social practice is to ask what it does for the participants of that practice. What needs does it answer to? What problems does it promise to solve? And how exactly does the practice function to meet the needs of the participants of that practice? Many philosophers have become attracted to the idea that we can uncover the function of a social practice by studying its structural origin.[10] The thought is that the general facts about ourselves and our environment which lead to the emergence of a social practice will

[10] This methodology is especially popular amongst philosophers of biology, e.g., Millikan (1984); Neander (1991); Kitcher (1993). See, Queloz (2020b) for a recent defence of genealogy as a method to reveal the social function of a social practice.

illuminate the needs the practice is supposed to address and explain how that practice might help to satisfy those needs. Since the best explanation for the emergence of a social practice is that the practice in question is disposed to respond to those needs of the participants of that practice. Thus, we can uncover the (current) function of a social practice by investigating its aetiology. This philosophical method of analysing a social phenomenon by investigating its aetiology is known as genealogy.

What does a genealogy involve? According to Williams (2002:22), a genealogy of a social practice is a narrative describing a way in which 'it came about, or could have come about, or might be imagined to have come about.' An investigation into the historical origins of a social practice can reveal the effect for which it was or could have been selected. Following Williams, we can distinguish between historical genealogies and imaginary genealogies. Let me illustrate the distinction with some examples.

Historical genealogy is a narrative of how a social practice came about based on historical facts. It describes the historical conditions under which a social practice developed over time. An accurate historical genealogy depicts the historical process and the evolvement of the social practice as it happened, supported by historical facts. Smyth's (2020) genealogical analysis of emancipatory values, i.e., values against oppression and dominance provides a recent example of a historical genealogy. Smyth argues that the emergence of emancipatory values is enabled by the cultural transition which places more emphasis on one's inner life; by the development in material and social resources and the growth of social connectedness; and triggered by the authoritarian terror and severe oppression from religion, state and the ruling economic class beginning from the 16th century. Leaving aside the question about accuracy, Smyth's genealogy is distinctively historical – he appeals to various historical facts such as the emergence of romantic novels, increased literacy, the growth of working classes in large cities, etc., to support his claim about the historical conditions under which emancipatory values developed.

However, using historical genealogy as a method to uncover the function of a social practice faces a few difficulties. First, in view of the complexity of a social practice and its causal connection to a diversity of events and forces, it is difficult to pin down the exact causal chain leading to the emergence of the social practice in question. Second, what justifies a particular selection of historical events as opposed to others? How do we determine whether the historical narrative in question is accurate? For example, is Smyth (2020) right in claiming that the emergence of sentimental novels plays a role in bringing about emancipatory values?

Smyth (2020:26) acknowledges that his account is tentative, and he expects readers to find fault with the historical narrative he gives, but stresses that the point is to show how illuminating a historical genealogy can be. This much should be granted. A tentative historical genealogy can in general enrich our understanding of a social phenomenon. However, a historically inaccurate genealogy would significantly undermine the justificatory force of the genealogical method in illuminating the function of the social practice in question.[11]

Imaginary genealogy, by contrast, avoids those difficulties. Imaginary genealogy is a fictional narrative which describes ways in which a practice could have come about, abstracting away from historical events and forces. Imaginary genealogies often involve state-of-nature stories, which describe a simplified, imaginary situation in which our ancestors lived. It reconstructs how a social practice could have developed under those initial conditions.[12] One influential example of imaginary genealogy in epistemology is Craig's (1990) genealogical analysis of the concept of knowledge. In Craig's state of nature, a small community of language-using human beings with unequal skills and talents were dependent on one another for their survival. Resources were scarce and threats from predators were constant. They had certain basic needs, including a need for cooperation which relies on obtaining true information effectively. Craig offers an account of how these basic needs could have led individuals to develop a prototypical concept of knowledge, *protoknowledge*, as a tool to flag good informants. He then describes how this prototypical concept is gradually objectivized – it loses its connection to the satisfaction of subjective needs and is developed into a concept with the properties associated with our concept of knowledge.[13]

Imaginary genealogy is abstract and idealised. Its epistemic force does not rely on the accuracy of its narrative, but on the plausibility of the counterfactual reasoning involved. Its main advantage is that, by constructing narratives in idealised conditions, it makes the alleged functional relation between the practice in question and the relevant needs stand out. Williams (2002:32-34) points out that imaginary genealogies are particularly apt at revealing the point or function of a phenomenon that is not obviously functional. Take Hume's (2000: 3.2.2) genealogy of the artificial virtue of 'justice' (understood as respect for property).

---

[11] My point here is not to reject the value of historical genealogies or deny the possibility of a responsible and accurate historical narrative of how a social practice emerged in history. Rather, I merely want to stress its difficulty and why the fictional narrative is to be preferred.

[12] State-of-nature stories feature prominently in political philosophy as a powerful philosophical tool to theorize how society and nation-states come into existence. See, for example, Hobbes (1994); Locke (1988); Rousseau (1992); Hume (2000); Rawls (1971); Nozick (1974).

[13] For more detailed reconstruction and further discussion of Craig's genealogy, see, Queloz (2019); Kusch (2009, 2011); Kusch & McKenna (2020).

Intuitively, we do not regard the virtue of justice or our disposition to respect for property as functional. But if Hume's analysis of the origin of justice is correct, it reveals that the amelioration of social conflict is the primary function of the virtue of justice. Craig's genealogy also illustrates this point. If his analysis is correct, then it reveals that the concept of knowledge has a certain social function, when we do not expect it to be so. Furthermore, Williams argues that, despite being fictional, the explanations imaginary genealogies offer are unmysterious. Imaginary genealogies explain what is functional in terms of human motivations and needs that are widely taken for granted. Hume's genealogy offers a game-theoretical reconstruction of the emergence of the virtue of justice based on a conception of humans as creatures that are self-interested and have limited sympathy.[14] Craig's analysis can also be read as a project of explaining the concept of knowledge in terms of the basic need for cooperation which requires obtaining true information effectively.

For these reasons, the genealogical analysis I offer in this section will rely mostly on fictional narratives and abstract away from specific historical conditions.[15]

3.4.2 Uncovering the Function of the T-practice

Using the method of genealogy, I will now turn to the task of uncovering the function of the T-practice. Drawing on the work done by Craig (1990) and Williams (2002), we can reconstruct how the T-practice could have emerged from epistemic state of nature. I will argue that the genealogy of the T-practice reveals that its social function is to facilitate knowledge production, maintenance, and social cooperation.

To see how the T-practice could have emerged, consider first the epistemic needs of early humans living in the state of nature (cf. Craig 1990: Chp.1; Williams 2002: Chp.3).[16] Early humans lived in small community with face-to-face interactions; they faced scarce resources and constant threats from predators. They relied on one another for survival. To meet the needs for food and shelter, they had to cooperate with each other. Cooperative social endeavours require information such as where the food was located, whether a predator was approaching, etc. In the state of nature, the need for survival gives rise to the

[14] For this reading of Hume, see Williams (2002: Chp.2).
[15] For further discussion on the methodology of genealogy, see Queloz (2020a).
[16] My description of the epistemic needs of early humans living in the state of nature is broadly in line with Craig (1990) and Williams (2002).

need for social cooperation, which in turn gives rise to the epistemic need for acquiring information about food, predators, shelter, etc.

Each member of the primitive society, regardless of their skills and talents, had what Williams (2002:42-43) calls *positional advantage*. Positional advantage is an epistemic advantage enjoyed by any epistemic agent simply in virtue of occupying a place at one time. Imagine an individual, Ada, living in the state of nature. Ada enjoys an epistemic advantage by occupying a spatial-temporal point which enables her to acquire information via observation available from that specific spatial-temporal point, e.g., whatever she perceives in her immediate environment.

It is worth noting, however, the positional advantage one enjoys is quite limited. When Ada's epistemic network is a single node, i.e., herself, the positional advantage she enjoys can only yield limited knowledge about the world (Figure 1). She can only acquire beliefs from what she manages to perceive in her immediate environment and from what she manages to infer on her own.

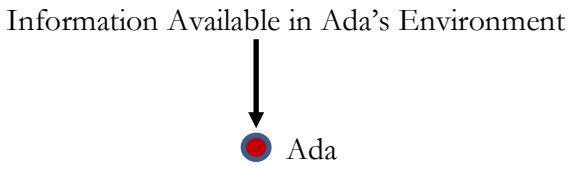


Figure 1: Epistemic Singleton

We can imagine that in the state of nature there were many epistemic singletons just like Ada, each enjoys some limited positional advantage. Positional advantage enables a primitive form of division of epistemic labour – by occupying a place at one time, each individual is in a position to acquire information by observation that may not be available to other members in their community. By sharing such information, the community as a whole can make the most out of each other's positional advantage to meet their collective epistemic needs for information about food, shelter, etc.

Imagine Ada, Ben, Chen and Dario, a small community of four individuals living in the state of nature. Suppose that each one of them, in virtue of their positional advantage, is in a position to acquire some important information about shelter, predators, food and water source. To acquire the vital set of information [shelter, food, water, predators] required for successful cooperation and their survival, they must engage in a division of epistemic labour and form an epistemic network such as the following:

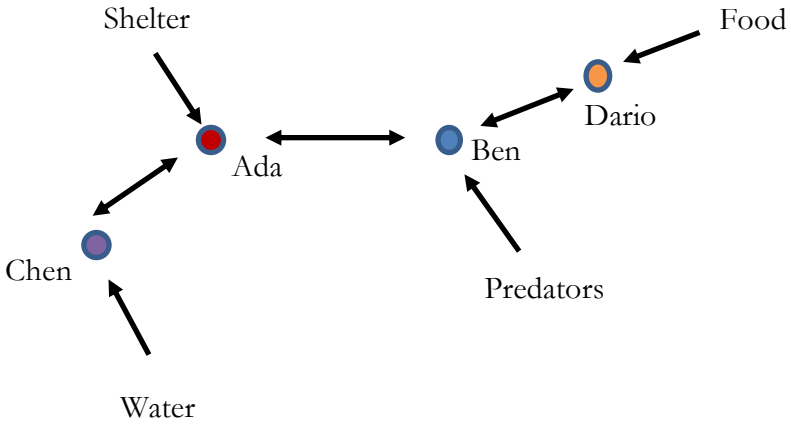


Figure 2: Epistemic Quad

Early epistemic communities like Ada's face a problem. To have a pool of information maximizing positional advantage, each individual must do their bit. They must acquire and maintain true beliefs from their environment and from their social network so that the vital set of information [shelter, food, water, predators] can become available to all members of the community.

Very likely, the T-practice could have emerged in response to this kind of challenge in primitive communities in the state of nature. Ada, Ben, Chen and Dario would have learned that to cooperate and survive, they must hold one another accountable with their belief-production and maintenance. It is in the interest of the whole community that everyone should do their best to acquire and retain true and only true beliefs. As the division of epistemic labour became more sophisticated and as more nodes were added to the epistemic network, the emergence of rule-governed patterns of interactive behaviour and attitude would have enabled communities to meet their collective epistemic needs and to ensure their survival. The genealogy of the T-practice in the state of nature suggests that it functions to pool information in a community to facilitate knowledge production, maintenance, and social cooperation.

What I said so far, I believe, gives a plausible account of how the T-practice could have emerged in early epistemic communities. But I can imagine someone raising the following two objections. First, granted that the brief genealogy offers some insight about the function of the T-practice in primitive societies. But why should the fictional narrative have any bearing on the *current* function of the T-practice in real, non-idealised conditions?

In a recent paper, Smyth (2017) makes a similar point in his criticism of the genealogical analysis of the function of morality developed in Joyce (2006) and Kitcher (2011). Smyth points out that functions are sensitive to environment. There is no such thing as 'intrinsic function'. The white fur of a polar bear has the function of camouflage only if its environment is white. But if that condition is altered, say, due to climate change, then the white fur may not have the same function. So, an entity E can perform its function in a system S only relative to a set of conditions. Call S* the current system in which E is said to have a certain function, Smyth formulates the requirement for a warranted inference from the genealogy of E to its current function as follows:

> Continuity: 'Given the set of enabling conditions that obtained in S when E developed, the same set of conditions obtains in S* and in the recent history of S*' (Smyth 2017:1132)

Smyth goes on to argue that the set of conditions which enables morality to perform its function in the state of nature (e.g., people living in a small community with face-to-face interactions, facing scarce resources and constant threats from predators and so on) no longer obtains in modern societies. The continuity condition is not satisfied. Hence, Smyth argues, the inference from the genealogy of morality to the claim about the current function of morality is significantly weakened.

Could one raise a similar objection regarding the T-practice? To begin with, some of the conditions which enable the T-practice to perform its function in the state of nature, e.g., people living in a small community with face-to-face interactions, facing scarce resources and constant threats from predators, enjoying positional advantage and so on, no longer obtain in modern societies. In the state of nature, our hunter-gatherer ancestors rely on the maximization of their positional advantage to generate a pool of information about food, shelter, etc. But in modern societies the positional advantage one enjoys in virtue of being at a place at one time plays limited role in guiding cooperative endeavours. Social cooperation in modern societies requires highly specialized knowledge. Obtaining such knowledge requires specific training and skills. The division of epistemic labour takes a different form. What is needed is expert input and an effective way to disseminate the relevant information to the relevant parties. Hence, it may be argued that the inference from the genealogy of the T-practice to the claim about the current function of the T-practice is significantly weakened due to the change of circumstances.

Smyth makes an important point that a function is relative to a set of conditions. But I do not think the exact same set of conditions must obtain for a warranted inference from the genealogy of E to its current function. For instance, it is not the case that the same climate and geographical conditions must obtain for the white fur of a polar bear to function as a camouflage. Suppose that polar bears are relocated to the Himalaya and survive. Even if the enabling conditions that obtained in the Arctic do not obtain in the Himalaya, the inference from the genealogy of the white fur of polar bears to its function is not undermined. What is required for a warranted inference, it seems to me, is a *structural continuity:* polar bears living in both environmental systems have a need to camouflage and the conditions in both systems enable their white fur to function as a camouflage. We should revise the Continuity condition as follows:

> Continuity*: Given the needs in response to which E emerged in S, the conditions in S* and in the recent history of S* must enable E to satisfy the same needs.

The move from the genealogy of the T-practice to the claim that it functions to facilitate knowledge production, maintenance and social cooperation does not violate Continuity*. Conditions in modern societies continue to enable the T-practice to carry out the same function as it first emerged in the state of nature. Let me explain.

As societies develop, of course, our specific epistemic needs change dramatically over time and the division of epistemic labour grows ever more sophisticated. In the state of nature, the kind of knowledge our hunter-gather ancestors needed for social cooperation was limited to mostly observational knowledge about the location of food, predators, etc. While in modern societies, the kind of knowledge we need for social cooperation is often systematic and highly specialised. But the expansion of our specific epistemic needs and the more sophisticated forms of division of epistemic labour do not violate Continuity*. There is a structural continuity between the state of nature and modern societies in that we continue to form epistemic networks to pool our epistemic resources and to engage in division of epistemic labour to produce and maintain knowledge and to facilitate social cooperation. Each individual epistemic agent in modern societies must also do their bit and hold one another accountable with their belief-production and maintenance. It is in the interest of the whole community that everyone should do their best to acquire and retain true and only true beliefs. The T-practice could be reasonably expected to persist in modern societies to meet our needs for knowledge production, maintenance, and social cooperation.

Another concern one might have for the genealogy of the T-practice is uniqueness. Suppose we define *E-practice* as a social practice of belief management governed by a norm of evidence and define *K-practice* as a social practice of belief management governed by a norm of knowledge. It seems that one could provide a similar genealogical analysis of the E-practice and the K-practice along the lines I have put forward for the T-practice. If so, it seems plausible to say that there may be a few alternative social practices that also arise in response to meet our needs for knowledge production, maintenance, and social cooperation.

But I do not think the observation that there are other social practices sharing the same social function as the T-practice raises any issue for my argument that the T-practice has the function of facilitating knowledge production, maintenance, and social cooperation. In fact, it would be surprising if the T-practice is the *only* social practice emerged over time to meet those needs, given how central they are to our survival and flourishing. Furthermore, I take the fact that similar analysis may be provided to account for the E-practice and the K-practice a booster to my view, for it shows the practice-based approach I developed in this chapter has the potential to not only account for the truth norm of belief, but also for the evidence and knowledge norm of belief. Indeed, a practice-based account of epistemic norms seems a promising project for further research.

In sum, the genealogy of the T-practice, i.e., how the T-practice could have emerged in early epistemic communities, reveals the current function of the T-practice, which is to facilitate knowledge production, maintenance, and social cooperation. I will now turn to the final moving part of my defence of the practice-based account of the truth norm: I will argue that the T-practice is a justified social practice and thus can properly ground the reason to believe p and to refrain from believing not-p when p is true.

3.4.3 Why the T-practice Is a Justified Social Practice

In Sect. 3.2.3, I emphasized that not all social practices can be justified by their social functions and many of them should be reformed or abolished altogether. I considered a few examples of unjustified social practices: practices that discriminate, oppress, or exploit the participants involved in those practices. I argued that a social practice cannot be justified by its social function if it causes substantial harm to its participants. Now, is the T-practice justified? Does it facilitate knowledge production, maintenance, and social cooperation at the cost of causing substantial harm to epistemic agents involved in that practice?

Recent discussions on epistemic injustices (e.g., Fricker 2007) and epistemic wrongdoings (e.g., Basu 2018, 2019) provide some grounds for concern here. For instance,

there is some evidence that facts about group membership affect how knowledge is produced and distributed in an epistemic community. If so, one might worry that the T-practice in its current form (and given the existing structural issues in our societies), may systematically favours members belong to the dominant groups and disfavour members belong the disadvantaged groups.[17] Worse still, there is some evidence that the T-practice may even cause harm to some epistemic agents involved in that practice. For example, believing certain things about a marginalized community, even when they are true, can often amplify an unjustified narrative about the community and put members of that community in harm's way.[18]

These cases, no doubt, raise critical issues. Indeed, I am quite sympathetic to the idea that the T-practice as it is being currently practiced is subject to reform and improvement. However, it is beyond the scope of my current project to address these issues directly here. Having said that, I believe there is a compelling case to be made for the T-practice. Even if the T-practice as it is being currently practiced is not flawless, there is overwhelming reason to think that it is justified, for even a partial breakdown of the T-practice would bring substantial harm to epistemic agents involved in that practice. In the remainder of this section, I will use George Orwell's *1984* as a thought experiment and suggest that the story brilliantly illustrates what a community would look like when the T-practice is in a partial breakdown. It would be marked by an impoverished epistemic environment and cause substantial harm to epistemic agents living in that community.

The story describes an imagined dystopian in the year 1984, when the world is divided into three totalitarian superpowers and is under perpetual war. The superpower Oceania is ruled by the Party under the dictatorship of Big Brother. The protagonist, Winston Smith, is a Party official and a secrete dissident. Winston works in the Ministry of Truth, which ironically controls the production and dissemination of misinformation and disinformation. In Oceania, the Party's machinations corrupt all standard epistemic resources. The Party controls all channels of news media – newspapers and radios reported 'news' according to the party lines: from economic data such as the quarterly output of pairs of boots (which always overfulfills quota) to the situations on the battlefield (which always glorifies the Oceanian army). The Party machinery also constantly destroys and revises historical records.

---

[17] It should be noted, however, it remains highly contentious whether it is the T-practice itself to be blamed here for epistemic injustices.

[18] I will return to this issue in Chapter 7. To give a quick preview: I think the right lesson to draw is not that the T-practice is an unjustified social practice. What these examples of epistemic wrongdoings suggest is that the T-practice-based reason may be disabled in some circumstances.

The Party frequently changes its alliance with one of the other two superpowers. Although Oceania was at war with Eastasia for years, once the Party declares that it is Eurasia who they had been fighting, all records concerning Oceania's relations with Eastasia and Eurasia are 'corrected'. And all traces of tampering are erased. There is no way to find out what happened by consulting historical records or documents. Furthermore, the entire society is under mass surveillance and ruled by terror. Thought Police is omnipresent through a surveillance device called telescreen and everyone is a potential victim of arbitrary arrest and torture. Most people are either brainwashed by propaganda or dare not to deviate from the Party line overtly. It is impossible to tell what others genuinely believe and no one is reliable. One's memory becomes unreliable due to long-term immersion in such corrupt epistemic environment. Winston's lover Julia, for example, does not remember that Oceania was at war with Eastasia and sincerely believes along the Party line that Oceania has been on friendly terms with them. Even sense perception becomes an unreliable epistemic source of belief regarding sensitive issues, because there is no other way to confirm what one takes oneself to see and hear.

When it concerns issues that are sensitive to the Party, such as politics, economy, warfare, etc., the epistemic environment in Oceania is so corrupt that none of the standard epistemic sources – testimony, memory and sense perception has a stable connection to truth. As such it is difficult for an Oceanian to correctly manage her beliefs even if she is committed to adhere to the truth norm of belief. Consequently, it is not surprising that Oceanians would persistently fail to exhibit patterns of interactive behaviour conforming to the truth norm of belief with respect to a large set of beliefs. A partial breakdown of the T-practice is marked by the impoverished epistemic environment caused by the Party's systematic manipulation.

Turning to another insight I want to draw from Orwell's *1984*. The daily running of the Party's machinations relies on Party members like Winston, who oversee the tampering of records, the production of propaganda and so on. So how could officials who have knowledge of the Party's manipulative tricks and manifold lies still fall victim of the Party's manipulation? Orwell offers a brilliant explanation by appealing to what he calls 'doublethink', a capacity to believe two contradictory claims simultaneously. It is through doublethink, a Party member can alter their memories and believe in things that they know to be false as long as the Party demands it. A new language – Newspeak is invented to reinforce people's ability to doublethink. In addition to the Party's active indoctrination of doublethink and the tools made available to facilitate doublethink, there is also a strong

incentive for a Party member to doublethink since this seems to be the only way to retain some sanity while running the machinations of the Party.

What Orwell describes as doublethink can be understood as a form of intentional self-deception: one deliberately brings about a belief that one knows or believes to be false (Bermúdez 2000; Talbott 1995; Davidson 1985). Orwell's story illustrates how it is possible for the Party to partially break down the T-practice even for those 'well-off' participants who do have some epistemic access to reality with respect to sensitive issues and are able to distinguish truth from falsity in a corrupt epistemic environment. It is the persistent and socially-embedded self-deception which makes it difficult for them to adhere to the truth norm.[19] Consequently, it is not surprising that even Party officials, the epistemic elites so to speak, also persistently fail to exhibit patterns of interactive behaviour conforming to the truth norm of belief with respect to a large set of beliefs. A partial breakdown of the T-practice is marked by pervasive self-deception induced by the Party's systematic manipulation.

Using Orwell's *1984* as an example, we see how the *T-practice* may be broken down in a community by a coercive, manipulative political force. When the *T-practice* is under attack, we pay a hefty price at both personal and societal level. Orwell draws a chilling picture for us:

> 'There will be no loyalty, except loyalty towards the Party. There will be no love, except the love of Big Brother. There will be no laughter, except the laugh of triumph over a defeated enemy. There will be no art, no literature, no science. When we are omnipotent we shall have no more need of science. There will be no distinction between beauty and ugliness. There will be no curiosity, no enjoyment of the process of life. All competing pleasures will be destroyed.' (Orwell 2003:306)

A person living in Oceania may have thoughts (e.g., thoughts along the Party lines), emotions (e.g., hatred towards the enemy and love of Big Bother) and relationships (e.g., they have families and colleagues). But none of them seems to mean anything anymore. They think, feel, and socialize only as dictated by the Party. They are neither curious nor creative. The living standards in Oceania are poor. Most people live in poverty. The productivity is low. Art and literature vanish because for those living in Oceania, there is nothing in themselves that seeks expression. People are entertained by pornography, cheap alcohol and state

[19] See Wei (2020), in which I discuss the social dimension of self-deception in more detail.

lotteries that never pay out. A life in Oceania is deprived of everything a human life is worth living for. At the personal level, one who loses contact with a large part of reality as a result of manipulation and self-deception is deprived of meaningful thoughts, emotions and relationships; at the societal level, we see a termination of knowledge production, a regress in technology, science and arts. Without the T-practice, a human life is hardly worth living.

Of course, Orwell's *1984* depicts a worst-case scenario unfolding in a totalitarian society. But a partial breakdown of the T-practice using the same kind of mechanisms described in the novel (e.g., propaganda, misinformation campaigns, etc.) is common in real life. Think about epistemic agents living in a religious cult, citizens in an authoritarian regime, or an online community of conspiracy theorists. Their epistemic situation is similar to someone living in the Oceania, in that none of them gain (hardly any) genuine knowledge from their social network.[20] The substantial harm caused by the partial breakdown of the T-practice is self-evident: cult members, oppressed citizens and victims of conspiracy theorists could not flourish in their impoverished epistemic environment. Orwell's dark tale of what a society would look like when the T-practice is under attack for political control is hardly an exaggeration. As one finds in Akhmatova's *Requiem*, a poem describing the pain and suffering of the Soviet people during the period of the Great Purge:

> 'In those years only the dead smiled,
> Glad to be at rest:
> And Leningrad city swayed like
> A needless appendix to its prisons.
> It was then that the railway-yards
> Were asylums of the mad;
> Short were the locomotives'
> Farewell songs.
> Stars of death stood
> Above us, and innocent Russia

[20] In a functioning liberal-democratic society, the epistemic environment is in fact in good standing since the standards of epistemic sources are well maintained by robust social institutions. Followers of religious cults and conspiracy theorists tend to adopt a sceptical attitude towards standard epistemic sources and as a result do not make use of those perfectly fine epistemic resources available to them. They take their wider epistemic environment to be corrupt unwarrantedly while at the same time over-relying on their immediate epistemic network of likeminded believers. For example, people who believe the conspiracy theory that vaccines cause autism express a strong distrust of the medical establishment. People who believe that the global warming is invented for ideological or financial reasons defying scientific consensus.

Writhed under bloodstained boots, and
Under the tyres of Black Marias.' (Akhmatova 1988:89)

Even death appears blissful in the horror of life.

In short, even a partial breakdown of the T-practice would cause substantial harm to epistemic agents at both individual and societal level. So, even if the T-practice as it is being currently practiced in our communities is not free of issues, overwhelmingly, the T-practice is a justified social practice and can properly ground the reason to believe p and to refrain from believing not-p when p is true. The practice-based account of the truth norm can meet the grounding challenge.

3.5 Conclusion

In this chapter, I argued for a practice-based account of the truth norm and showed that it can meet the grounding challenge. The grounding challenge, recall, concerns whether the reason to believe truly and only truly can be properly grounded in some reason grounding fact. I developed and defended a practice-based account, according to which, the authority of the truth norm is grounded in a justified social practice, i.e., the T-practice. The reason to believe p and to refrain from believing not-p when p is true is practice-based.

I first motivated the practice-based account by contrasting it favourably with the evaluative and the constitutive account of the truth norm. There are three motivations for the practice-based account: (1) it considers the social aspect of belief; (2) it avoids the wrong-kind-of-reason problem; and (3) it avoids the problem of limited 'authority'. I then argued that the T-practice is a social practice satisfying all three earmarks of social practices. Next, I contended that the T-practice functions to facilitate knowledge production, maintenance, and social production. Given that its social function is indispensable to the survival and flourishing of humankind and given that even a partial breakdown of the T-practice would cause substantial harm to its participants, I conclude that the T-practice is a justified social practice. The T-practice can ground the reason to believe p and to refrain from believing not-p when p is true. Thus, the practice-based account offers an attractive response to the grounding challenge.

Part II. Can the Truth Norm Guide?

Summary of Part I and Preview of Part II

In Part I, I addressed the grounding challenge. On the reason-based normative framework, if the truth norm of belief is authoritative, then the reason to believe truly (correctly) and to refrain from believing falsely (incorrectly) must be properly grounded. In Chapter 2, I considered and criticized existing attempts to ground the reason on value-based, desire-based and constitutive reason grounding facts. In Chapter 3, I developed and defended a practice-based account of the truth norm. I showed that the reason to believe truly and only truly can be properly grounded on the practice-based fact that if p were true believing p and refraining from believing not-p would constitute a justified social practice, i.e., the T-practice.

In Part II, I turn to the guidance challenge. Norms are the sort of things that can provide guidance. On the reason-based normative framework, if the truth norm of belief can provide guidance, then it must be possible for us to respond to the T-practice-based reason. In Chapter 4, I will first present in more detail the No Guidance argument and explain why some philosophers think that the truth norm cannot guide our belief-formation. I will argue that the existing response to the No Guidance argument is inadequate in light of emerging empirical evidence about our system of belief-formation. I will motivate an alternative response to the No Guidance argument, and present, in rough outline, a reason-responsive account of epistemic guidance (REG). To be guided by the truth norm to form a belief p, is for one to respond to the T-practice-based reason in forming that belief. In Chapter 5, I will identify a problem for applying REG on the truth norm of belief. To solve the problem, I will develop a novel account of what it takes to respond to the T-practice-based reason. My central idea is that believing p in response to the T-practice-based reason is a *mental action*. For one to respond to the T-practice-based reason is for one to either manifest the relevant disposition to respond to the T-practice-based reason or manifest a distinctive phenomenology of the mental action of believing something in response to the T-practice-based reason.

4. The Guidance Challenge

Norms are the sort of things that can provide guidance. So, if the truth norm of belief is a genuine norm, it must be capable of guiding our belief-formation. In recent discussion, however, it has been argued that on a standard *inferentialist* model of guidance, the truth norm cannot guide belief-formation.

In this chapter, I will argue that the inferentialist model is not the only way to understand guidance. I will propose an alternative, reason-responsive model of guidance, according to which to be guided by an epistemic norm to form a belief is to respond to the relevant normative reasons. On the reason-based framework, to be guided by the truth norm to form a belief p is to act for the T-practice-based reason to believe p when p is true. The plan is as follows. I will first provide an overview of the state of the debate. I will begin by presenting Glüer and Wikforss' *No Guidance* argument and briefly consider a popular response to the No Guidance argument which maintains the inferentialist model of epistemic guidance (Sect. 4.1). I will argue that this response is inadequate in light of emerging empirical evidence about our system of belief-formation (Sect. 4.2). I will motivate an alternative response to No Guidance which rejects the inferentialist account of epistemic guidance (Sect. 4.3). I will then present, in rough outline, a reason-responsive model (REG) of epistemic guidance and show that the truth norm can provide guidance on our belief-formation (Sect.4.4). I will further develop and defend a version of REG in the next chapter.

4.1 The State of the Debate

4.1.1 The No Guidance Argument

Glüer and Wikforss (2009, 2010, 2013) argue that the truth norm of belief is not normative because the truth norm cannot guide our belief-formation. They start their argument by sketching what they claim to be an intuitive picture of guidance. They claim that any generic guiding rule has the following form:

(R) Do X when C.

To be guided by R, S must first form a *belief* as to whether the trigger condition C is met (see, e.g., 2010:758, 2013:83). Call this the Doxastic Constraint. Second, S's doing X must be *in virtue of* her acceptance of R: 'to be guided by a norm or rule in our performances intuitively requires that R influences or motivates or provides reasons for these performances.'

(2010:758). In other words, if S's doing X is guided by R, then S's belief that C obtains, and her acceptance of R must motivate S's doing X (and justify S's doing X if R is a norm). Call this the Reason Constraint.

On this intuitive picture of guidance, we can make sense of how a rule like 'buy low and sell high' can guide. To be guided by that rule is for S to first come to have a belief as to whether the price is low or high, and for that belief and S's acceptance of the rule to explain and rationalize S's selling or buying things. However, Glüer and Wikforss argue, we cannot make sense of how the truth norm of the form:

(T) Believe p when p is true,

can guide our belief-formation on this intuitive picture of guidance, since the Reason Constraint cannot be satisfied. They argue:

> 'First, the norm cannot influence belief-formation, since in order to be guided by [the truth norm] one *already* had to form a belief as to whether p. The very question we wanted guidance on – whether to believe that p – needs to have been answered before any such guidance can be received. Thus, [the truth norm] always comes 'too late', as it were, to do the guiding work it is supposed to do. Second, [the truth norm] never makes a difference in the following sense, either: it never "tells me" or "provides me with a reason" to believe anything other than what I have come to believe anyway.' (Glüer and Wikforss 2013:84, their emphasis; see also, 2010:759)

One natural way to make sense of their argument, is to interpret Glüer and Wikforss as endorsing an inferentialist model of guidance.[1] The inferentialist model of guidance is attractive because it captures what it is for the subject to act *in virtue of* being guided by R as opposed to merely happening to do X when C. The intuitive idea is that the 'in virtue of' relation can be cashed out in terms of S's performing an inference. To be guided by the truth norm, on the inferentialist model of guidance, is to infer in accordance with the truth norm.

If one tries to infer in accordance with the truth norm, we see that the input belief required by the Doxastic Constraint is the very same belief that one seeks guidance for. Guidance from a norm is supposed to be normative: your response should gain some

[1] See, for instance, Boghossian (2008); Wright (1989, 2007).

normative support in virtue of being guided. On the inferentialist model of guidance, the fact that you make a correct inference in accordance with the norm in question should provide you with some justification for doing X. But a problem arises when one tries to make an inference in accordance with the truth norm. Inferring a belief in accordance with the truth norm does not seem to provide justification for one's belief since it seems to involve inferring a belief p on the basis of the very same belief that p. But most beliefs – if not all beliefs – cannot justify themselves in such inferences. So, one may argue, inferring in accordance with the truth norm cannot provide normative support for believing p.

Moreover, it seems that inferring a belief in accordance with the truth norm cannot be the reason for which you come to have that belief. If I ask you why you believe that the Covid-19 vaccine is effective, and your answer is that you are guided by the truth norm, i.e., you infer the belief after you formed the belief that the Covid-19 vaccine is effective, you do not really explain to me the reason for which you believe that the Covid-19 vaccine is effective. For being guided by the truth norm does not seem to be the motivating reason for which you form the belief that the Covid-19 vaccine is effective. It seems that the Reason Constraint cannot be satisfied in the case of following the truth norm. For any S, p, S's inferring p in accordance with the truth norm provides neither motivation nor justification for S's believing p. We can reconstruct No Guidance as follows:

(1) A genuine norm is essentially capable of guiding – one must be able to act in response to that norm.
(2) For any S, p, if S's belief p is guided by a norm, then S must infer the belief p in accordance with S's representation of that norm, which provides both motivation and justification for S's believing p.
(3) For any S, p, it is impossible for S to believe p by making an inference in accordance with the truth norm, which provides both motivation and justification for S's believing p.
(4) Therefore, for any S, p, it is impossible for S's belief p to be guided by the truth norm. (2&3)
(5) Therefore, the truth norm is not genuinely normative. (1&4)

In this chapter, I will discuss two possible responses to No Guidance. I will argue that rejecting premise (3) is inadequate in light of emerging evidence about our belief-forming system. I will then motivate an alternative, more demanding response to No Guidance, which is to reject premise (2).

4.1.2 Rejecting Premise (3)

One popular response to the No Guidance argument is to reject premise (3) and show that making an inference in accordance with the truth norm can provide both motivation and justification for one's belief (e.g., Steglich-Petersen 2010, 2013).

Consider first why we might think inferring in accordance with the truth norm can provide normative support for one's belief. We can make a distinction between reasons and enabling conditions under which one can respond to a reason. Your belief that the price is low is an enabling condition for you to act on that reason. But that enabling condition, i.e., your belief, itself is not the normative reason for which you act under the guidance of 'buy low and sell high'. The reason in favour of buying things for which you act is *the fact that* the price is low. Likewise, the input belief about whether p is true merely plays an enabling role for one to apply the truth norm. As an enabling condition, the input belief p is not itself the reason in favour of believing p for which one believes p under the guidance of the truth norm. Importantly, we should not conflate the conditions under which the truth norm recommends beliefs with the psychological states one must have in order to apply the norm. The condition under which the truth norm recommends belief is the fact p, not the belief p. With this clarification, one may argue, to infer in accordance with the truth norm is *not* to form a belief p on the basis of the very same belief p, but to infer a belief from a fact. Thus, making an inference in accordance with the truth norm can provide normative support for believing p.

Consider next why we might think inferring in accordance with the truth norm can provide a motivating reason for belief. One might think that Glüer and Wikforss have a too narrow understanding of what it takes for a norm to influence or motivate one's belief-formation. To see this, we can imagine S adopting a rule of pleasure, which tells her to believe p just in case believing p is pleasant. To be guided by the pleasure rule, S would first try to determine whether believing p is pleasant and make an inference about whether to believe p in accordance with that rule. Had S instead been guided by the truth norm, she would have directed her efforts to determine whether p is true. Now, suppose that p is true but believing p is unpleasant. If S were to be guided by the truth norm, she would believe p. If S were to be guided by the pleasure rule, she would (all else equal) refrain from believing p. By contrasting the two cases, we can see that S's belief can be influenced and motivated by her making an inference in accordance with the truth norm, since S could have ended up with a different belief had she adopted a different rule such as the rule of pleasure. Whether one

makes an inference in accordance with the truth norm or the pleasure rule does make a difference as to whether one would have that belief or not.

These considerations raise serious doubts about premise (3). However, in the next section, I will argue that emerging empirical evidence on our system of belief-formation poses a problem for this general line of response. Since the problem is not simply that making an inference in accordance with the truth norm cannot provide motivation or justification for one's belief, but rather, forming a belief by making an inference in accordance with the truth norm does not seem to be something we can (or usually) do given the kind of belief-forming system we have, empirically speaking.

4.2 Spinozan System of Belief-formation

Empirical studies of belief in fields such as delusion, self-deception, heuristics, biases, cognitive dissonance, and conspiracy theories have been growing at a rapid speed in the past decades.[2] The emerging scientific picture of our belief-forming system is one that is fast, automatic and subject to many psychological and social factors often out of our direct control. It is a picture at odds with the traditional Cartesian picture held by many epistemologists, which is the backbone of the inferentialist model of epistemic guidance. On the Cartesian picture of belief-formation, one first entertains a proposition p and then either arrives at the belief p or rejects it.[3] It is during this evaluative process one may seek guidance from epistemic norms, which tell us whether to form the belief p or to reject it. The emerging scientific picture of belief, however, aligns much better with the so-called Spinozan account of belief-formation (Gilbert 1991; Gilbert, Tafarodi, and Malone 1993; Egan 2008; Mandelbaum 2014). On the Spinozan account, one forms beliefs more or less automatically when representational contents are made available through cognitive mechanisms such as perception and they are not products of a norm-invoking, effortful inference at the person-level. Only a relatively small number of beliefs will go through a further evaluative process during which some of the beliefs may be rejected, provided that they are subject to an unfavourable verdict.[4] In this section, I will first argue that our belief-forming system is likely

---

[2] For an overview, see Bortolotti (2009) on delusion, Von Hippel & Trivers (2011) on self-deception, Gilovich et al. (2002) on heuristics and biases, and Douglas et al. (2019) on conspiracy theories.

[3] Rejecting a proposition p includes two possible doxastic states: one may either suspend judgment about p or one might arrive at the belief not-p.

[4] The Spinozan account of belief-formation can be seen as part of dual-process/dual-system theories of the mind, which is default-interventionist in structure (Evans & Stanovich 2013; Kahneman & Frederick 2002). According to default-interventionist theories, Type 1 processing (e.g., automatic and

to be Spinozan for a number of reasons. I will then explain in more detail why this poses a problem for the common response to the No Guidance argument we considered above.

To begin with, it would seem that from a purely evolutionary point of view, it is more likely that our belief-forming system is Spinozan. Beliefs play important action-guiding roles when paired with the relevant desires, and it is essential for us to have efficient belief-forming system that can help us navigating in a fast-changing world. Given our limited cognitive resources, it is a good thing to have a belief-forming system which can quickly and reliably add information to our stock of beliefs without taking extra cognitive efforts. If we had to undergo an evaluative process to turn the information we acquired through, say, perception into belief, then we would not be able to take up perceptual beliefs as efficiently and reliably, especially when we are put under significant cognitive load, e.g., when we are multi-tasking, as we often do. Furthermore, going through an evaluative process takes up additional cognitive resources. So, one may expect that a well-adapted effective mental economy is one where we apply evaluative processes selectively to only a small number of beliefs, e.g., beliefs flagged as suspicious or important.

Of course, there is a risk of feeding a sizeable number of false beliefs into our mental economy with the Spinozan system of belief-formation in place, since beliefs are added to our stock of beliefs without first being subject to evaluation and scrutiny. However, it is plausible that, the cost of failing to add information to our belief stock, or delaying the uptake, may well be higher than the cost of the occasional uptake of false beliefs. When I see that a car is coming towards me at a high speed, there is a much higher chance to survive if I act on the unevaluated perceptual belief and run to safety, than having a belief-forming system which takes time to first evaluate the perceptual experience before delivering it as a belief which then guides me to run to safety.

In addition to this admittedly somewhat suggestive evolutionary consideration, there are important empirical results that provide evidence for the Spinozan hypothesis (Gilbert 1991; Gilbert et al. 1993; Skurnik et al. 2005; Masip et al. 2006; Hasson et al. 2005; Unkelbach 2007). For example, Gilbert et al. (1993) designed six experiments which placed one group of participants under a cognitive load condition –- a disabling performance constraint –- to make it harder for the participants to go through an evaluative process and compared these with a control group of participants who were not placed under a cognitive load. Broadly

---

reflexive processing) generates default responses on which subsequent reflective Type 2 processing (e.g., reflective and controlled processing) may or may not intervene.

speaking, a Cartesian account of belief-formation would predict that participants under a cognitive load condition are less likely to arrive at a doxastic state compared to the control group and that the cognitive load will affect the acquisition and the rejection of a belief equally. In contrast, a Spinozan account would predict that additional cognitive load will disproportionally affect a subject's rejection of a belief rather than the acquisition of a belief. In each of the six experiments conducted by Gilbert et al. (1993), researchers found that participants who were under a cognitive load were much worse at rejecting falsehoods and arrived at more false beliefs than those who were not under a cognitive load, but that both groups were almost equally good at arriving at true beliefs, inconsistent with the prediction of the Cartesian account.[5]

Furthermore, on the Cartesian account, what underlies our belief-formation is a controlled and reflective evaluative process.[6] However, it is widely acknowledged that what often underlies our belief-formation in decision-making in a wide range of contexts is *not* a controlled or reflective evaluative process but heuristics and biases, which allow us to make quick and effective judgments with relatively low cognitive costs.[7] Ample empirical evidence on heuristics and biases casts doubt on the Cartesian account. The Spinozan account, by contrast, is easily compatible with the evidence that heuristics and biases underlie our belief-formation. On the Spinozan account, one forms beliefs more or less automatically when representational contents are made available through various cognitive mechanisms, which can be susceptible to the influence of heuristics and biases.

Together with heuristics and biases, the Spinozan hypothesis offers a better explanation of a much broader scope of phenomena such as self-deception, conspiratorial thinking, and the easy uptake of misinformation when one's epistemic environment is impoverished. In

---

[5] It is debatable, however, whether affirming a statement p is a good indicator that the participant holds the belief that p, and the extent to which the structure of the task affects the result. For further criticism, see Street and Richardson (2015). Nevertheless, these results provide some *prima facie* evidence against the Cartesian account.

[6] This kind of process is often referred to as Type 2 processing by psychologists. See footnote 12.

[7] There is a wide range of evidence suggesting that there are various heuristics underlying our belief-formation (Kahneman & Tversky 1974), such as the availability heuristics (mental processes used to make judgment about probability based on the ease with which the relevant content can be brought into mind), representativeness heuristics (mental processes used to make judgment based on stereotypes and prototypes), anchoring and adjustment heuristics (mental processes used to make judgment depending heavily on an 'anchor', e.g., an initial piece of information offered or an antecedent belief). There are also a wide range of biases underlying our belief-formation, such as confirmation bias, a tendency to form beliefs that support or confirm the beliefs one already has (Klayman & Ha 1987); partisan bias, a tendency to form beliefs based on one's partisan identity (Van Bavel & Pereira 2018); self-enhancing bias, a tendency to form self-enhancing beliefs (Alicke & Sedikides 2009), and many more.

paradigm cases of self-deception, for example, self-deceptive subjects form beliefs insensitive and irresponsive to the evidence available to them even if they do subscribe to some epistemic norms and are in a position to apply them correctly (e.g., they have the cognitive capacities to make inferences and possess the relevant evidence).

Let me illustrate this point with an example. Consider Lena, whose beliefs are by and large sensitive and responsive to evidence. However, she is self-deceptive when it comes to her career prospects. She believes that her debut novel is an enormous success, that she will receive a fellowship at a prestigious university, and so on, even though she has plenty of evidence suggesting the contrary. On the Cartesian picture, it is puzzling why a self-deceptive subject like Lena, who is otherwise a responsible epistemic agent, would persistently fail to evaluate a sizeable number of propositions concerning her career prospects. If an evaluative process underlies each belief-formation, why would Lena come to believe that her novel is an enormous success and that she will receive a fellowship, when she has plenty of evidence that the novel has not been well received, given that she subscribes to epistemic norms and is in a position to apply them correctly? Likewise, it is puzzling, why many epistemic agents, who are otherwise responsible epistemic agents, would *persistently* fail to correctly evaluate many propositions concerning conspiracy theories and other misinformation.

Now, of course, there are various viable psychological and philosophical explanations as to why an otherwise responsible epistemic agent may persistently fail to evaluate her beliefs correctly. My point here is to highlight that these cases would not appear to be puzzling in the first place if we recognize that our belief-forming system is Spinozan. These phenomena are natural consequences of our belief-forming system being Spinozan. Let me explain.

On the Spinozan account, evaluation does not occur prior to belief-formation. Self-deceptive, conspiratorial and misinformed beliefs are added to the subject's belief stock in automatic, reflexive processes susceptible to the influence of various heuristics and biases.[8] These self-deceptive and conspiratorial beliefs are likely to be 'protected' from undergoing a further evaluative process due to various social and psychological factors, e.g., it is painful for Lena to believe that her debut is a total failure. Beliefs in conspiracy theories often help alienated individuals to regain a sense of control in a fast-changing world.[9] If self-deceptive and conspiratorial beliefs do not typically undergo an evaluative process in the Spinozan

[8] The Spinozan picture fits well with theories of self-deception, according to which self-deception consists of biased belief-formation influenced by mental states such as anxiety (Barnes 1997), desire (Mele 2001) or pretense (Wei 2020).

[9] For an overview of the various factors underlying conspiratorial thinking, see, for example, Douglas et al. (2019).

belief-forming system of those epistemic agents who are under certain social and psychological conditions, that explains why those epistemic agents who are otherwise responsible epistemic agents can end up with a large number of self-deceptive or conspiratorial beliefs. It also explains why the uptake of misinformation can be easy and widespread among epistemic agents whose epistemic environment is impoverished. On the Spinozan picture, beliefs are added to our belief stock from our information ecosystem in automatic, reflexive processes, and only a relatively small number of them will go through a further evaluative process. So, for those epistemic agents living in an impoverished epistemic environment filled with false information, there is no surprise that they might form a sizable number of false beliefs despite most of them are responsible epistemic agents and subscribe to epistemic norms.

Of course, none of the considerations I presented above are conclusive, but taken together, they make a compelling case that our belief-forming system is likely to be broadly Spinozan. That is, a large part of our belief-formation is unlikely to undergo a Cartesian evaluative process prior to the production of beliefs. Our belief-forming system is likely to be default-interventionist in structure (Kahneman & Frederick, 2002; Evans and Stanovich 2013).

How does this emerging picture of our belief-forming system bear upon the debate on the No Guidance argument? In the previous section we have seen that the dispute between normativists and anti-normativists has focused on the third premise. Normativists argue that inferring in accordance with the truth norm can provide both motivation and justification for one's belief-formation. But the emerging evidence on our belief-forming system suggests that this line of response is inadequate. For the issue normativists must address is not simply that if one were to infer in accordance with the truth norm whether that can provide both motivation and justification for one's belief. Rather, the issue seems to be that given our belief-forming system is Spinozian in character, how belief-formation, quite generally, can be the product of making an *inference* in accordance with the truth norm. This, I believe, is a serious and underdiscussed challenge for normativists who want to reject premise three and show that the truth norm can provide guidance on the inferentialist model.

Let us unpack this challenge more carefully and consider two possibilities. First, if our *entire* belief-forming system is Spinozan, then no belief can be formed as a product of making an inference in accordance with the truth norm. Rather, they are formed in automatic, reflexive processes when representational contents are made available through cognitive mechanisms such as perception. On this picture, only a small number of beliefs undergo a

further evaluative process, during which an effortful, norm-invoking inference at person-level may occur. So, if our entire belief-forming system is Spinozan, no belief-*formation* can be guided by the truth norm, or indeed by any norm whatsoever, although some of our belief-*revision* may be guided by epistemic norms.

Indeed, this consideration has motivated some to endorse the view that all epistemic norms are norms of belief-*revision* (e.g., Helton 2020). We cannot assess this position here, but I think it is fair to say that most philosophers, including the critics of the truth norm, reject global scepticism about epistemic norms of belief-*formation*. The main takeaway point, in any case, is that no epistemic norm can guide belief-*formation* on the inferentialist model of guidance if our entire belief-forming system is Spinozan, which is bad news for normativists who accept the inferentialist model of guidance.

But of course, whether our *entire* system of belief-formation is Spinozan is a debatable empirical matter. This leads us to the second possibility, given the evidence, it may well be that our belief-forming system is *largely* but not entirely Spinozan. Now, even if only a large part of our belief-forming system is Spinozan, we are still forced to conclude that a sizeable proportion of our belief-*formation* cannot be guided by the truth norm (or any other epistemic norms) on the inferentialist model of guidance. Only a small number of beliefs formed through a Cartesian evaluative process can be guided by the truth norm (i.e., if they are the products of inferring in accordance with the truth norm in cases of doxastic deliberation). However, this is also a bad result for normativists who would like to use the truth norm to do the sort of philosophical heavy lifting we sketched in Chapter 1. If our belief-forming system is largely Spinozan, then it is not the case that for any S, p, S's belief p can be guided by the truth norm, as far as belief-formation is concerned. The truth norm can at best be a norm governing a small number of beliefs formed in the Cartesian system and a small number of beliefs revised in the Spinozan system. Most beliefs are not subject to or guided by a norm of truth. Normativists would therefore have to give up on the general claim that belief is subject to and guided by a norm of truth.

Thus, rejecting premise (3) while maintaining an inferentialist model of epistemic guidance is not a satisfying response for normativists in light of emerging empirical evidence about our belief-forming system.

4.3 Motivating an Alternative Response

An alternative, more demanding response to the No Guidance argument is to reject premise (2). It is a more demanding response since, first, the inferentialist model of guidance is

popular among philosophers. Rejecting this intuitive model of guidance seems an unappealing and costly move. Second, normativists who adopt this response would have to provide a non-inferentialist model of epistemic guidance and show that the truth norm can guide on that model. Neither task is easy. In the remainder of this chapter, I want to make some headway on both fronts.

Despite the popularity of the inferentialist model of guidance, recent work on inference and normative theories has casted some doubts on this model. One is the problem of vicious regress raised by Boghossian (2008). The problem arises because inference itself is often understood as a kind of transition *guided by rules of inference*. This rule-guided conception of inference in combination with the inferentialist model of guidance leads to vicious regress. To illustrate, consider Boghossian's 'Email Rule':

> (Email Rule) 'Answer any email that calls for an answer immediately upon receipt!' (Boghossian 2008:481)

On the inferentialist model of guidance, being guided by the Email Rule involves an inference in accordance with the Email Rule: I must first form a belief that the condition for applying that rule is met (i.e., that I received an email that calls for an answer), and then draw the conclusion with an intention to do what the rule requires me to perform and act accordingly.

There is a problem, however, since inference itself is *guided* by a rule of inference. Making an inference in accordance with a rule is to be guided by a rule of inference like the following:

> (Inference Rule) If 'Do X when C' and C, do X.

Now if the inferentialist model of guidance is correct, being guided by the Email Rule involves making an inference in accordance with the Email Rule, which in turn involves being guided by the Inference Rule. If being guided by the Inference Rule itself requires yet another inference guided by the relevant rules of inference involved, then on the inferentialist model of guidance, making that inference would again require making another inference… Therefore, the combination of the inferentialist model of guidance and a rule-guided conception of guidance would lead us to an infinite vicious regress. To avoid the regress, something has to go: either inferring in accordance with a rule is *not* for that rule to guide the agent's inferential behaviour, or the inferentialist model of guidance is incorrect.

The problem of vicious regress has motivated some philosophers to reject the rule-guided conception of inference and propose that instead we should understand inference, for example, as a distinctive kind of response to some informational states that produces a

conclusion (Siegel 2019); or as a transition guided by rules built into the architecture of cognitive systems (Quilty-Dunn & Mandelbaum 2018); or as responses regulated by some inferential pattern (Boghossian & Wright (ms)).

It is beyond the scope of this chapter to assess these alternative conceptions of inference. The point here is that the problem of vicious regress puts pressure on the rule-guided conception of inference *as well as* the inferentialist model of guidance. While much of the existing responses to the problem of vicious regress focus on revising our conception of inference, it is worth highlighting that by moving away from the rule-guided conception of inference we in effect also move away from the inferentialist model of guidance found in the No Guidance argument. According to the latter, being guided by a norm involves an effortful, norm-invoking inference which provides both motivation and justification for one's belief-formation.

Another potential difficulty for the inferentialist model of guidance is what I will call *the problem of unworthy guidance*. It is a familiar thought in ethics that one's action can be morally correct while lacking moral worth. The most famous example is Kant's shopkeeper who treats his customers fairly but only out of self-interest. The shopkeeper does the morally right thing but his action has no moral worth: he has no genuine concern or respect for moral laws (Kant 4:397). One natural explanation for the lack of moral worth of the shopkeeper's action is in terms of normative guidance. The shopkeeper's action lacks moral worth because his action is not guided by moral norms but by self-interests. An analogous notion of epistemic worth can be applied to beliefs: one's beliefs can be correct while lacking *epistemic worth* because the epistemic agent might believe the right thing while having no concern for truth or evidence (Whiting 2020). A norm-guided belief has an epistemic worth that is lacking in a belief not guided by a norm.

On the inferentialist model, to be successfully guided by a norm is to make a correct inference in accordance with that norm. However, the inferentialist model faces a challenge: there are cases where the subject makes the correct inference in accordance with the relevant norm but nevertheless her response lacks the kind of worth that norm-guided responses deserve.

Consider Gendo who recently made a substantial donation to the World Food Programme. Suppose that he did so by making a correct inference in accordance with the norm of beneficence. However, Gendo does not think there is any moral reason to be beneficent and he believes that all moral talk is nonsense. For Gendo, a 'norm' of beneficence is just like any other societal rule that he follows when doing so is expedient for him to

navigate through the world. It seems that Gendo's action lacks moral worth because he acts out of self-interests and shows no concern for morality. On the inferentialist model of guidance, however, Gendo's action is guided by the norm of beneficence. Since responses guided by a moral norm has moral worth, it follows that Gendo's action has moral worth. But this contradicts the intuitive verdict that Gendo's action lacks moral worth.

Analogous problems arise in epistemic cases. It is easy to imagine that Gendo disregards the truth norm too. Gendo does not think there is any epistemic reason to believe the truth. For Gendo, the truth norm, too, is just like any other societal rule he follows when doing so is expedient for him to navigate through the world. Suppose that Gendo forms his beliefs by making correct inferences in accordance with the truth norm (leaving aside the controversy regarding whether this is possible). It seems that Gendo's beliefs lack epistemic worth because he does not have a genuine concern for truth. On the inferentialist model of guidance, however, Gendo's beliefs are guided by the truth norm and thereby have epistemic worth. But this contradicts the intuitive verdict that Gendo's beliefs lack epistemic worth.

It seems then, the inferentialist model of guidance allows for *unworthy guidance*, i.e., cases where one's responses are guided by norms but are not worthy. To avoid the problem of unworthy guidance, something has to go: either the intuitive understanding of moral and epistemic worth in terms of normative guidance is mistaken, or the inferentialist model of guidance is incorrect. Again, it is beyond the scope of this chapter to assess alternative accounts of moral and epistemic worth. The point here, is that the problem of unworthy guidance puts pressure on the inferentialist model of guidance for those who accept the intuitive understanding of moral and epistemic worth in terms of normative guidance.[10]

Of course, neither the problem of vicious regress nor the problem of unworthy guidance is a knock-down argument against premise (2), but they show two things. First, proponents of the No Guidance argument cannot simply take the inferentialist model of guidance for granted. In fact, a lot more needs to be said to shore up the inferentialist conception of epistemic guidance. Second, our brief discussion shows that despite the popularity of the inferentialist model of guidance, it has been challenged in the wider

---

[10] One possible response is to insist that one cannot be said to be making a correct inference in accordance with a norm if that person does not recognize the normative force of the norm. Since Gendo does not recognize the beneficence norm or the truth norm as genuinely normative, he cannot be said to be making correct inferences. So Gendo's responses are not norm-guided on the inferentialist model and thereby lack moral and epistemic worth. But this response seems *ad hoc*: there is no independent reasons why performing an inference requires one to appreciate the normative status of the inference rule. If we want to maintain the idea that norm-guided responses are creditworthy, it seems that we are forced to reject the inferentialist model of guidance.

literature. Thus, rejecting premise (2) is not a move as undesirable as it might first appear to some normativists. It provides motivation for normativists to explore alternatives to the inferentialist model of guidance, which leads us to the second challenge facing normativists who reject premise (2).

4.4 The Reason-responsive Model of Epistemic Guidance

The main challenge facing normativists who adopt the more demanding response to the No Guidance is to provide an alternative, non-inferentialist model of guidance. A natural move, as I suggested in Chapter 1, is to develop an account of epistemic guidance on a reason-based framework: to be guided by an epistemic norm is for one to respond to the relevant epistemic reasons in forming one's beliefs. In what follows, I will outline an account of epistemic guidance in terms of reason-responsiveness and highlight how the truth norm can provide guidance on this alternative model in ways that avoid the difficulties facing the inferentialist account.

The basic idea is simple: for one's belief-*formation* to be guided by the truth norm is for one to form a belief p in response to the reason that p is true.[11] On this model, to be guided by the truth norm is fully compatible with the empirical evidence suggesting that our belief-forming system is largely, if not entirely Spinozan. Let us begin our discussion with a characterization of the reason-responsive model of epistemic guidance (REG):

> (REG) For any S, p, S's forming a belief p is guided by an epistemic norm according to which R is a reason to believe p, just in case S forms the belief p in response to the reason R when R is available to S.

On this model, then, to be guided by the truth norm is for S to form a belief in response to the reason that p is true when the fact that p is available to S. In general, a reason R to believe a proposition is available to S only if S has some epistemic access to R, e.g., S is in a position to know R through perception, intuition, testimony, memory, reasoning or other sources of knowing a fact.[12] So for example, the fact that the pavement is wet is a reason available to S

---

[11] I focus on belief-formation because this is what causes the most trouble for the normativists given that our belief-forming system is Spinozan.

[12] Just what exactly this access amounts to is controversial. Other popular candidates include, for example, knowing the fact (e.g., Williamson, 2000) and having reflective access to the fact (e.g., Audi, 1993).

only if S is in a position to know that the pavement is wet, e.g., by seeing that the pavement is wet, remembering that the pavement is wet and so on.

To form a belief in response to a reason R is for S to manifest a capacity of reason-responsiveness in believing p. The capacity of reason-responsiveness is typically understood in dispositionalist terms. The basic idea of the dispositionalist account is that to believe p for a reason R is for S to manifest her capacity to use R as a reason to believe p, which consists of some of S's epistemic dispositions.[13] A disposition, recall, is a function mapping some input conditions to the corresponding output conditions. An epistemic disposition is a function mapping certain input conditions to S's doxastic attitudes. According to the dispositionalist account of reason-responsiveness, S believes p for a reason R just in case S's believing p manifests some of S's epistemic dispositions mapping some input conditions about R to output conditions about S's believing p. To fix ideas, I will follow Lord's (2018:138-139) formulation of the dispositionalist account of responding to a normative reason:

> (DRR) S forms a belief in response to a reason R just in case S's believing p manifests a disposition to believe p when R constitutes a reason to believe p and to revise the belief p when R ceases to be a reason to believe p.[14]

There are two ways in which R may cease to be a reason for believing p. First, since reasons are factive, R ceases to be a reason when it goes from obtaining to not obtaining. For example, the consideration that the pavement is wet ceases to be reason in favour of believing that it is raining when the pavement is not in fact wet. Second, R ceases to be a reason when it is disabled and thus provides no normative support for believing p. For example, the fact that the cup appears red to me ceases to be a reason for me to believe that the cup is red when the lighting condition in my room is not normal. If one has the relevant reason-responsive disposition to believe p when R, then the following counterfactual is usually true: in the absence of overriding reasons, S would not believe p if R were to cease to be a reason

---

[13] Limited by space, I cannot engage with non-dispositionalist views of reason-responsiveness such as *rationalism,* according to which reason-responsiveness consists of acts of deliberation (Barry 2007; Chan 1995; Korsgaard 2009), or *psychologism*, according to which reason-responsiveness consists of having a belief-desire pair which causally explains why one responds in the way one does in the presence of the relevant reasons (Davidson 1970; Arpaly & Schroeder 2012, 2013, 2015). I think dispositionalist views in general are preferable to non-dispositionalist views for they offer an attractive solution to the problem of deviant causal chains. See, e.g., Lord (2018: Chp.5).

[14] For dispositionalist views in the vicinity, see, Mantel (2017:562-3); Kiesewetter's (2017:164). In Chapter 5, I will offer a novel account of what it is to respond to the T-practice-based reason.

for believing p and S would believe p if R were to obtain, provided that S has epistemic access to R. So for example, given that I would not believe that the cup is red if the fact that the cup appears red ceases to be a reason for me to believe that the cup is red (e.g., when the reason is disabled) and I would believe that the cup is red when the cup appears red is a reason for me to believe that (e.g., when the reason is not disabled), my belief manifests my capacity to respond to that particular perceptual evidence as a reason for belief.

Applying REG and DRR to the truth norm, we can say that, to be guided by the truth norm is for S to manifest the relevant reason-responsive disposition in forming the belief p, i.e., the disposition to believe p when the fact that p is true constitutes a reason for believing p and to revise the belief p when that reason ceases to obtain.

Of course, much more should be said about REG and there are potential objections to be answered, a task I will turn to in the next chapter. But for our present purposes, it suffices to motivate REG by contrasting it favourably with the inferentialist model of guidance. REG is a promising account of epistemic guidance because it satisfies four criteria for a viable model of epistemic guidance implicit in our previous discussion:

(1) A viable model must be able to explain what it is for a belief to be guided such that guidance is more than mere conformance;
(2) A viable model should be able to explain how beliefs can be guided in Spinozan system of belief-formation;
(3) A viable model should not be viciously circular, i.e., it must not account for guidance by appealing to conditions that are themselves explained in terms of guidance;
(4) A viable model should be able to account for the epistemic worth of guided beliefs and avoid unworthy guidance.

REG satisfies all four criteria whereas the inferentialist model, as we have seen, has trouble meeting many of them. First, according to REG, to be guided by an epistemic norm is for S to form a belief in response to the relevant epistemic reasons in a way that manifests S's reason-responsive capacity. Merely forming a belief p in the presence of those epistemic reasons without manifesting the relevant epistemic dispositions, on REG, does not count as being guided by that epistemic norm.

Second, REG enables us to explain how an epistemic norm can guide belief-formation in a way compatible with both Cartesian and Spinozan account of our system of belief-formation. The relevant reason-responsive dispositions, in principle, can be manifested in

both a controlled, reflective process where one first entertains a proposition p and then either arrives at the belief p or rejects it, as well as in an automatic, reflexive process where beliefs are added to one's stock of beliefs without undergoing evaluation. Every belief can be guided by epistemic norms on REG, even if our entire system of belief-formation is Spinozan.

Third, on REG, guidance is explained in terms of reason-responsiveness. The capacity of reason-responsiveness, according to the dispositionalist account I presented above, consists of dispositions to believe p when R constitutes a reason to believe p and to revise the belief p when R ceases to be a reason to believe that p. The capacity of reason-responsiveness is not understood in terms of guidance. REG does not run into a vicious regress.

Fourth, REG offers a natural explanation as to why guided beliefs are epistemically worthy. A guided belief has epistemic worth because it is a belief formed in response to the relevant epistemic reasons available to S such that S manifests her capacity of reason-responsiveness. Recall that Gendo has no genuine concern for truth. For him, all normative talk is nonsense and he does not treat truth as a reason for belief. He believes whatever that is expedient for him to navigate through the world. The inferentialist model has trouble explaining why Gendo's belief formed by inferring in accordance with the truth norm lacks epistemic worth, for guided beliefs should have epistemic worth. REG, by contrast, can deliver the right verdict. Since Gendo's belief-formation does not manifest his capacity to respond to truth *as a reason* for belief, Gendo's belief is not guided by the truth norm and thus his belief lacks epistemic worth. So, REG avoids the problem of unworthy guidance.

We are now able to say how the truth norm can guide our belief-formation on REG in ways that avoid the difficulty facing the inferentialist account. Regardless whether our actual belief-forming system is largely or entirely Spinozan, for any S, p, it is possible for S to form a belief in a way that manifests her reason-responsive capacity to believe p when p is true constitutes a reason for believing p and to revise the belief p when that reason ceases to obtain. There is no principled reason as to any epistemic agent would not be capable of manifesting such capacity in Spinozan system of belief-formation. So, on REG, for any S, p, it is possible for S to form a belief p following the guidance of the truth norm. The truth norm can guide our belief-formation on REG.

4.5 Conclusion

I have argued in this chapter that rejecting premise (3) is an inadequate response to the No Guidance argument in light of emerging empirical evidence about our system of belief-

formation. I then motivated a more demanding response to No Guidance argument, and showed that, on REG, for any belief, under either Spinozan or Cartesian belief-forming system, the truth norm can provide guidance and thus can be genuinely normative. REG provides a welcome alternative to the inferentialist model of guidance and offers a promising way to understand how our belief-formation can be guided by the truth norm of belief.

At this point, one might argue that by adopting REG, I am simply pushing the guidance challenge one step back: I must now defend the idea that the truth norm of belief can provide guidance on REG understood in dispositionalist terms. But I think this is a step forward in the right direction. In the next chapter, I will address a problem for applying REG in combination with DRR on the truth norm of belief. I will show that the T-practice-based reason is a reason we can respond to and thus we can be guided by the truth norm.

5. Believing for the T-practice-based Reason

In the previous chapter, I argued that the common response to the No Guidance argument is inadequate considering the emerging evidence on our belief-forming system. I motivated an alternative response and argued that we can meet the guidance challenge by adopting a reason-responsive model of epistemic guidance, REG. Understood in dispositionalist terms, REG satisfies all four criteria for a viable model of epistemic guidance. The truth norm of belief can provide guidance: for one to be guided by the truth norm is just for one to manifest dispositions to respond to the T-practice-based reason when that reason is available.

In this chapter, I will respond to a problem for the proposal I have put forward so far. The problem arises because, as I will explain in more detail, it seems that a novice of the T-practice cannot be guided by the truth norm of belief on REG understood in dispositionalist terms. This is highly problematic, for a novice is most in need of guidance. It seems then, there is a remaining puzzle about the guiding capacity of the truth norm. And the normative status of the truth norm seems dubious if it cannot provide guidance to novices of the T-practice.

The aim of this chapter is to provide a solution to this Novice Problem and in doing so I will develop a novel *hybrid dispositionalist/phenomenological* account of what it takes to respond to the T-practice-based reason. My central idea is that believing p in response to the T-practice-based reason is a *mental action*. For one to respond to the T-practice-based is for one to either manifest the relevant dispositions to respond to the T-practice-based reason or manifest a distinctive phenomenology of the mental action of believing something for the T-practice-based reason. This account will allow us to make sense of how a novice can be guided by the truth norm even though she has not yet possessed the relevant reason-responsive dispositions to respond to the T-practice-based reason.

The plan is as follows. I will first introduce and diagnose the Novice Problem (Sect. 5.1). I will defend the claim that believing something for reasons is a mental action (Sect 5.2). Drawing on Anscombe's (1957) idea that one has non-observational knowledge of one's own action if one acts for reasons, I will propose an account of the phenomenology of believing something for the T-practice-based reason (Sect. 5.3). Finally, I will put these pieces together and articulate a novel hybrid dispositionalist/phenomenological account of what it takes to respond to the T-practice-based reason. I will show how my account can deal with the Novice Problem and provide a satisfactory response to the guidance challenge (Sect. 5.4).

5.1 Guiding the Novice

5.1.1 The Novice Problem

Let us begin by rehearsing the proposal I have put forward so far. On REG, to be guided by the truth norm is for S to form a belief in response to the T-practice-based reason that p is true when that reason is available to S. The T-practice-based reason to believe a proposition is available to S only if S has some epistemic access to that reason through perception, intuition, testimony, memory, reasoning, or other sources of knowing a fact. To form a belief in response to the T-practice-based reason to believe p when p is true is for S to manifest a reason-responsive capacity in believing p.[1]

Following Lord (2018), the capacity of reason-responsiveness is understood in *dispositionalist* terms (DRR): to form a belief in response to a reason R is for S to manifest a disposition to believe p when R constitutes a reason to believe p and to revise the belief p when R ceases to be a reason to believe p. So, on DRR, to form a belief in response to the T-practice-based reason that p is true is for S to manifest a disposition to believe p when the T-practice-based constitutes a reason for believing p and to revise the belief p when the T-practice-based reason ceases to be a reason for believing p.

Note that most proponents of DRR take the reason-responsive disposition to be a *primitive* notion which resists further reductive analysis. However, it is widely accepted that the manifestation of a disposition is typically understood as involving modal robustness. If one manifests the relevant disposition to believe p in response to the T-practice-based reason, then the following counterfactual is usually true: in the absence of overriding reasons, S would not believe p if the T-practice-based reason were to cease to be a reason for believing p and S would believe p if the T-practice-based reason were to obtain, provided that S has epistemic access to that T-practice-based reason. The falsity of the counterfactual usually indicates that S's believing p does not manifest the relevant reason-responsive disposition and is not guided by the truth norm of belief. Although the reason-responsive dispositions are not analysed in counterfactual terms, counterfactuals can be a good guide to whether one manifests the relevant dispositions or not (e.g., Lord 2018:138 fn.24).[2]

---

[1] For the ease of presentation, I will focus our discussion on *forming a belief* in response to the T-practice-based reason. But the same dispositionalist account can be offered to *refraining from forming a belief* in response to the T-practice-based reason.

[2] There are two familiar reasons why a counterfactual account of reason-responsive is false. First, on the standard Lewisian view, counterfactual conditionals are vacuously true when the antecedent clause of that counterfactual conditional describes an impossibility (Lewis 1973). This causes a familiar problem, since some T-practice-based reasons are necessary truths. When a T-practice-based

However, an intuitive thought about guidance poses a problem for reason-responsive model of epistemic guidance understood in dispositionalist terms (e.g., REG & DRR). It is natural to think that when one first begins to engage in a social practice, one goes through a learning process. One learns about what it is to follow the guidance of the relevant rules governing that practice including what the correct rules governing that practice are and how to apply them in a situation. Let us define a novice as an agent who has recently been introduced to and adopted a social practice. At the point of adoption, the novice accepts the rules governing the practice, but she is prone to make mistakes. Since a novice is someone who has only recently been introduced to and adopted a social practice, it is natural to assume that she is in the early stage of a learning process regarding how to follow the guidance of the rules governing that practice. It is expected that the novice is prone to make mistakes about how to apply the rules in a situation.

Consider a child who has recently been introduced to and adopted the T-practice. She learns that it is correct to believe p just in case p is true. Plausibly, she may not be particularly good at telling whether a proposition p is true and constitutes a reason to believe p or discerning what the correct application of the truth norm is in a situation. It takes her time and experience to learn. It seems impossible for a novice, at the early stage of her learning process, to fully possess the capacity to use the T-practice-based reason as a reason for belief. She is incapable of manifesting dispositions to respond to the T-practice-based reason to believe p when p is true. It follows then, on the account I have put forward so far, a novice of the T-practice cannot be guided by the truth norm of belief.

This is an unpalatable result – the truth norm of belief cannot provide guidance if one is a novice of the T-practice. But a novice to the T-practice is most in need of guidance! We can put the Novice Problem in terms of an incompatible triad:

---

reason is a necessary truth, the relevant counterfactual 'S would not believe p if the T-practice-based reason had not obtained' is vacuously true, regardless of whether S manifests the relevant dispositions. Therefore, when the T-practice-based reason is a necessary truth, the dispositionalist cannot appeal to the truth-value of the counterfactual as a guide to whether S manifests the relevant reason-responsive dispositions, since the counterfactual is vacuously true. Second, in cases of overdetermination, the falsity of the relevant counterfactual does not always indicate that one fails to manifest the relevant dispositions to believe p in response to the T-practice-based reason. This is because when there is another reason R* for which S believes p, S would still believe p (for R*) even if the T-practice-based reason to believe p were to cease to obtain. This causes a problem, since the T-practice-based reason is usually not the only reason for which one believes p. One usually believes p for various evidential reasons. Therefore, when S's believing p is overdetermined, i.e., when there is more than one reason for which S believes p, the falsity of the counterfactual does not indicate the S fails to manifest the relevant dispositions to respond to the T-practice-based reason.

(REG & DRR)	For any S, p, S's forming a belief p is guided by the truth norm just in case S manifests the relevant dispositions to respond to the T-practice-based reason when the T-practice-based reason is available to S.
(NOVICE)	A novice of the T-practice cannot manifest the relevant dispositions to respond to the T-practice-based reason.
(GUIDANCE)	A novice of the T-practice can be guided by the truth norm of belief.

So, it seems that I have not solved the guidance challenge after all! If the truth norm of belief cannot provide guidance to novices, it is dubious that the truth norm is a genuine norm capable of providing guidance.

5.1.2 Moving Towards a Hybrid Dispositionalist/Phenomenological Account

One might think that the Novice Problem is simply a case in which the counterfactual fails to be a guide to whether the novice manifests the relevant reason-responsive dispositions. All it shows is that the falsity of the relevant counterfactual does not necessarily indicate that the novice fails to manifest the relevant dispositions to respond to the T-practice-based reason.[3] But this does not undermine DRR, after all, according to DRR, the reason-responsive dispositions are primitive. The counterfactual is never meant to be a necessary and sufficient condition for the manifestation of the relevant reason-responsive dispositions.

But the problem runs deeper, I think. The novice case is not simply a case where the counterfactual is not a good guide to whether the novice manifests dispositions to respond to the T-practice-based reason. Rather, the problem arises because by the very definition of a novice, a novice is incapable of manifesting the relevant dispositions to respond to the T-practice-based reason to believe p for she is yet to acquire those dispositions. So, the Novice Problem forces us to accept that the manifestation of the relevant reason-responsive dispositions is not a necessary condition for one to respond to the T-practice-based reason, though it might be a sufficient condition. We should revise our account of reason-

[3] That is, in the absence of overriding reasons, a novice would not believe p if the T-practice-based reason were to cease to be a reason for believing p and the novice would believe p if the T-practice-based reason were to obtain, provided that the novice has epistemic access to that T-practice-based reason.

responsiveness. To solve the Novice Problem, I propose the following hybrid dispositionalist/phenomenological account:

> (HRR) S believes p for a reason R just in case S's believing p either (1) manifests dispositions to believe p when R constitutes a reason to believe p and to revise the belief p when R ceases to be a reason to believe p, or (2) manifests a distinctive phenomenology of believing p in response to the T-practice-based reason as a mental action.

On HRR, to form a belief in response to the T-practice-based reason is for S's believing p either to manifest the relevant reason-responsive dispositions or to manifest a distinctive phenomenology. In the remainder of this chapter, I will develop and defend the hybrid dispositionalist/phenomenological account of what it is to respond to the T-practice-based reason, focusing on the phenomenological component. I will argue that REG understood in the hybrid terms can deal with the novice problem.

5.2 Believing Something for Reasons as a Mental Action

The central idea I aim to develop and defend is that believing p in response to the T-practice-based reason is a *mental action* with a distinctive phenomenology.[4] The manifestation of the distinctive phenomenology is a sufficient condition for one to respond to the T-practice-based reason. Although a novice of the T-practice is unable to manifest the relevant reason-responsive dispositions in her belief-formation, she can manifest the distinctive phenomenology if her belief-formation is guided by the truth norm. So, a novice can be guided by the truth norm on REG, understood in the hybrid dispositionalist/phenomenological terms.

Let me begin my defence of the hybrid account by addressing an immediate objection which says that since belief is a mental state and mental states cannot be actions, believing something in response to reasons cannot be a mental action. According to this line of objection, mental states are static – they obtain over and throughout intervals of time. Believing p is not an aspect of the mind that, once obtains, would undergo changes, or unfold over time. Mental actions, by contrast, are dynamic mental events – they are a kind of mental

[4] The idea that believing can be something we do in our capacity as epistemic agent is fairly uncontroversial (e.g., Korsgaard 2009; Peacocke 2008; Boyle 2011), though some are more sceptical (e.g., Setiya 2013).

happenings that occur at times or unfold over periods of time. Although mental states and mental actions may both have temporal extension, they do not have the same temporal character. Since mental states and mental events are ontologically distinct, and mental actions are mental events, believing something in response to reasons cannot be a mental action. Thus, one might argue, appealing to the phenomenology of S's believing p in response to reasons as a *mental action* is a non-starter.

However, this objection fails to make a distinction between *belief* and *believing*. Every belief that obtains over a period of time is marked by the occurrence of three mental events: the acquisition of that belief, the process of sustaining that belief and if it ceases to exist, the revision of that belief. While a *belief p* is a mental state, *believing p*, e.g., acquiring or sustaining that belief, is a mental event. Therefore, although S's belief p is a mental state and hence cannot be a mental action, the mental event of S's *believing p* can be a mental action.

Naturally, not all mental events are mental actions and hence not every case of believing something is a mental action. However, the intuitive support for the claim that believing something *for reasons* is a kind of mental action seems strong. To see this, let us draw an analogy between acting for reasons and believing for reasons. It is fairly uncontroversial that things we do for reasons are actions performed in our capacity as an agent. Consider the famous opening vignette in Davidson's (1980) seminal essay on agency:

> 'This morning I was awakened by the sound of someone practising the violin. I dozed a bit, then got up, washed, shaved, dressed, and went downstairs, turning off a light in the hall as I passed. I poured myself some coffee, stumbling on the edge of the dining room rug, and spilled my coffee fumbling for the New York Times.' (Davidson 1980:43)

There is a clear contrast between things I do in my capacity as an agent (e.g., getting up, washing, shaving, dressing, going downstairs, turning off a light, pouring coffee, fumbling for the New York Times) and things that I do as a result of something happening to me (e.g., stumbling on the edge of the rug, spilling coffee). In contrast to the things I do that fall within the latter category where I am merely the subject of some bodily movements ascribed to me, the things I do that fall within the former category are actions.

This distinction can be extended to things I do *mentally*. There is a similar contrast between things I do mentally in my capacity as an agent (e.g., deliberating, calculating, listening to Bach's Goldberg Variations) and things I do mentally as a result of something happening to me (e.g., hearing the sound of a heavy rain shower, being startled by the sight

of a lightning strike, wandering off in a day dream, remembering a childhood episode triggered by the taste of madeleine). In contrast to the things I do mentally that fall within the latter category where I am merely the subject of some mental events ascribed to me, things I do mentally that fall within the former category are mental actions.

Further, and as Peacocke's (2008:255) observes, performing a mental action such as calculating has an active phenomenology which is lacking in a mental event such as daydreaming where the subject seems to be passively related to the event befalling her. Korsgaard (2009) also emphasizes the active aspect of acting and believing for reasons:

> '…the space of reflective distance presents us with both the possibility and the necessity of exerting a kind of control over our beliefs and actions that the other animals do not have. We are, or can be, *active, self-directing*, with respect to our beliefs and actions to a greater extent than the other animals are, for we can accept or reject the grounds of belief and action that perception and desire offer to us. We can *actively participate* in giving shape both to the conception of the world in light of which we act and to the motives on the basis of which we act – and ultimately, in both ways, in giving shape to ourselves.' (Korsgaard 2009:32, my emphasis)

It is notoriously difficult to articulate a principle by which we can distinguish actions from mere doings. For our purposes, however, there is no need to draw such a strict conceptual line here. It suffices to say that the contrast between the two kinds of doings sketched above holds and it is intuitive that things we do for reasons belong to the first kind of doings, i.e., they are actions.

To strengthen the case that things we do for reasons are best characterised as actions, consider the following argument:[5]

> (1) The question 'for what reasons does S φ?' has application when S's φ-ing is a response to reasons.
>
> (2) The question 'for what reasons does S φ?' has no application when S's φ-ing is not done in her capacity as an agent but merely because of something befalling S.

[5] This argument draws on, but is independent from, Anscombe's (1957) claim about the applicability of a certain 'why' question.

(3) Therefore, if S φ-s for reasons, then S's φ-ing cannot be done merely because of something befalling S, but is something done in her capacity as an agent, i.e., φ-ing for reasons is an action.

Premise (1) seems uncontroversial. If S's φ-ing is a response to reasons, then the question 'for what reasons does S φ?' has a meaningful answer. We will focus our discussion on premise (2).

Imagine that you ask me for what reasons I behaved angrily towards my father and I answer, 'I don't know' or 'no reason'. My failure to provide an answer does not, however, imply that the question has no application. According to Anscombe (1957), there are three kinds of cases where the question 'for what reasons does S φ?' is not applicable.

First, the question has no application when S does not know what she is doing. What would be left of our description of S's φ-ing if S lacks knowledge of what she is doing? Imagine S, who is walking towards a supermarket. Under one plausible description, S is walking towards a supermarket to buy food. Now, if we subtract the fact that S knows that she is walking towards a supermarket, what is left of our description of S's walking towards a supermarket to buy food? It seems that we cannot plausibly describe S's walking in the street as an action subject to the question 'for what reasons is S walking towards the supermarket?' The question asking for S's reasons for walking towards the supermarket has no application.[6]

Second, the question has no application when S knows what she is doing but only on the basis of observation. Imagine that S is emptying a bottle of liquid. Suppose that it is only on the basis of observing the liquid coming out of the bottle that S knows that she is emptying a bottle of whisky. It seems that we cannot plausibly describe S's emptying that bottle of whisky as an action subject to the question 'for what reasons is S emptying that bottle of whisky?' The question asking for S's reasons for emptying that bottle of whisky has no application.

---

[6] One well-known criticism to this line of thought is that the claim that S must *know* what she is doing if the question asking for S's reasons has application is too strong. Consider a famous example from Davidson (1970:92). A carbon-copier intends to make ten legible carbon copies but does not know that he is making ten legible carbon copies since he is not sure whether he will succeed, yet it seems that we can plausibly describes the carbon-copier's making carbon copies as an action subject the question asking for the carbon-copier's reasons for making ten carbon copies. One way to respond is to make a distinction between knowing that one is φ-ing and knowing that one is φ-ing successfully. Although the carbon-copier fails to know whether he has succeeded or will succeed in making ten legible carbon copies, he does know that he is making ten legible carbon copies. So, the case is not a counterexample.

Third, the question asking for S's φ-ing has no application when the only plausible explanation of S's φ-ing is a mental cause:

> 'A mental cause is what someone would describe if he were asked the specific question: what produced this action or thought or feeling on your part: what did you see or hear or feel, or what images or ideas cropped up in your mind and led to it?... I want to distinguish it from the ordinary senses of "motive" and "intention".' (Anscombe 1957:17-18)

Suppose that the only plausible explanation of S's smashing the window is a mental cause, e.g., that S suffers from an intermittent explosive disorder. It seems that we cannot plausibly describe S's smashing the window as an action subject to the question 'for what reasons is S smashing the window?' The question asking for S's reasons for smashing the window has no application.

Now, it seems that, things S does as a result of something befalling her nicely match Anscombe's description of the three kinds of cases where the question 'for what reasons does S φ?' has no application. Consider Davidson's vignette again. I can only know that I stumbled and spilled my coffee based on observation and the only plausible explanation of my behaviour is some mental cause such as that I was distracted. It seems that we cannot plausibly describe my stumbling over the edge of a rug and spilling my coffee as actions subject to the question 'for what reasons did you stumble and spill your coffee?' The question asking for my reasons for behaving in this way has no application. So, if what Anscombe account about cases 'for what reasons does S φ?' has no application is plausible and cases in which S φ-s because of something befalling her fits Anscombe's account, there is a compelling reason to think that the question 'for what reasons does S φ?' has no application when S φ-s because of something befalling her. Thus, if S φ-s for reasons, S's φ-ing cannot merely result from something befalling her: φ-ing for reasons is an action.

To sum up, in this section I defend the claim that believing something in response to reasons is a mental action. First, I dismissed the objection according to which believing p is a mental state and hence cannot be a mental action. Second, I showed that there is a clear contrast between two kinds of doings – believing for reasons seems to clearly fall within the category of mental actions. Finally, I argued that believing something for reasons is a mental action from the applicability of the question asking for one's reasons for φ-ing. In the next section, I will turn to the *phenomenology* of believing something in response to reasons as a mental action.

5.3 The Phenomenology of Believing Something for Reasons

In our discussion of the applicability of the question which asks for S's reasons for φ-ing, we touched upon the Anscombean idea that S has non-observational knowledge of her φ-ing if S φ-s for reasons, then. In this section, I will focus on the phenomenological aspect of having non-observational knowledge of one's own action. I will suggest that S's believing p is a response to the T-practice-based reason to believe p, if S manifests an action-awareness which is distinctively non-observational. I will identify two phenomenological features of that non-observational action-awareness: (1) it is as of something that has a certain direction – moving towards the aim of believing something true and (2) it is as of something that is subject to some *self-imposed* constraints in the service of achieving that aim.

Let us begin with the idea that the phenomenology of acting for reasons involves a non-observational awareness of one's action.[7] Consider a few common examples. When I raise my fingers over the keyboard to write down a thought, I am aware of my typing, even if I do not perceive the movement of my fingers or their points of contact with the keyboard. Similarly, my awareness of many of my everyday actions such as going to a supermarket, lighting up an incense, fetching a book, etc., does not seem to rely on the perception of my bodily movements.

Of course, this is not to say that I cannot become aware of my actions through perception.[8] Perception usually plays an active role in enabling me to carry out an action successfully. For example, perception enables me to make various adjustment so that I can successfully type the word 'justice'. It helps to shape the way I act as my action unfolds. To say that the action-awareness is non-observational when I am acting for reasons, is therefore *not* to deny the existence of observational action-awareness, but rather, to emphasize that I am not a mere spectator of my action. There is a distinctively practical route from which I can gain such action-awareness which is not available to others. To become aware of what I

---

[7] The account I offer below focuses on the phenomenology. It is consistent with, among others, Peacocke's (2008) influential account of action-awareness. But it does not involve any commitment to such account. According to Peacocke, action-awareness is representational, has first-personal, present-tensed content of the form 'I am doing such-and-such now'. It is caused by an event of trying and it makes available demonstrative ways of thinking of actions.

[8] Many have made the point that observation should be distinguished from perception. Non-observational knowledge is not equivalent to non-perceptual knowledge. Perception can play a part in acquiring non-observational knowledge (Gibbons 2010; O'Shaughnessy 1963; Pickard 2004; Schwenkler 2011).

am doing is not a journey of discovery, e.g., I will not be surprised to find out that I am typing the word 'justice'.

Now, since believing for reasons is a mental action, believing for reasons also involves non-observational action-awareness. But what is it for one to be non-observationally aware of her believing p?

Observational awareness of one's bodily action is an awareness of one's bodily action on the basis of observing the bodily events/states involved. Analogously, we might think that observational awareness of one's mental action is an awareness of one's mental action on the basis of observing the mental events/states involved.[9] So we might say that one is non-observationally aware of her believing p if one is aware of her believing p without observing the mental events/states involved.

However, there is a problem. Although observational awareness of bodily events/states is naturally understood as an awareness delivered by sensory perception, it is less clear what the counterpart is in the mental case. One way to make sense of observational awareness of one's mental states/events is to think of it as an awareness delivered by a quasi-perceptual mechanism. Let us use the term 'introspective perception' to refer to whatever causal mechanism that does the job.[10] I will not commit myself to any particular theory of introspective perception and leave it to the reader to decide whether introspective perception is even possible, and if it is possible, how to best understand the nature of introspective perception.[11] Nothing in particular will hinge on this here, since our claim concerns *non*-observational awareness of the mental events/states involved in one's believing p. Talk of observational awareness based on introspective perception is useful to us only in so far as it helps us to think more clearly – in an analogous way to the case of bodily action – about non-

---

[9] However, there is an additional sense in which one's awareness of one's mental action can be observational: since one's bodily behaviour could be the manifestation of a mental state/event, it is possible for one to be aware of one's mental action via the perception of one's bodily behaviour.

[10] How to understand introspective perception is central to perceptual (quasi-perceptual) accounts of introspection. A minimal perceptual (quasi-perceptual) account satisfies the following conditions: (1) perceptual beliefs are caused by objects of perception and (2) objects perceived are independent of the possibility of perceiving them.

[11] Shoemaker (1988, 1994) famously argued against the possibility of introspective perception on the ground that it entails the possibility of a phenomenon he called 'self-blindness'. Since it is impossible to be self-blind, introspective perception is impossible. For those who think introspective perception is possible, they differ on how it should be understood. Introspective perception has been understood as the operation of a mental self-scanning or self-monitoring process (e.g., Armstrong 1968:324; Lycan 1996; Nichols and Stich 2003); or as having a higher-order belief about one's first-order mental states (e.g., Soteriou 2013); or as acquaintance (e.g., Gertler 2001), as the interpretation of a complex sub-personal mechanism (e.g., Carruthers 2011); or as constitutive and implicit self-awareness (e.g., Zahavi 2005).

observational awareness of one's mental action. The idea, then, is that for S to be non-observationally aware of her believing p is for her to be aware of her believing p *without introspectively perceiving* the mental events/states involved in S's believing p.

To illustrate, consider the mental action of believing p through a process of deliberation. A process of deliberation can involve various mental states, such as beliefs, desires, fears, etc., and various mental events, such as perceiving, visualizing, making an inner speech, entertaining a proposition p, etc. Suppose that I am deliberating about whether it is a good idea to purchase a new sofa. Suppose that the following mental states and events are involved in the process: a desire to have a new sofa; a visualization of the sofa in my living room; an inner speech on the cost and benefit of buying a new sofa. It seems that I can be aware of what I am doing mentally even if I do not introspectively perceive the desire I have, the imagery I was visualizing, or the inner speech I made. There is a practical route by which I can be aware of my believing that it is an excellent idea to buy a new sofa, simply by making up my mind that it is an excellent idea to buy a new sofa.

Again, this is not to say that I cannot become aware of my believing p by introspectively perceiving the various mental state/events involved, or that introspective perception plays no role in obtaining the non-observational awareness of my believing p. Introspective perception of various mental states/events can play an active role in enabling me to successfully carry out a mental action as it unfolds. For example, introspective perception of what I am visualizing – whether I am visualizing the sofa and my living room with the right details can help to shape the way I continue to visualize things as my deliberation about whether to purchase a new sofa unfolds. Neither do I deny the possibility of gaining awareness of my mental action through observing my bodily behaviour, i.e., seeing myself pacing in my living room, hearing myself thinking out loud and in excitement and declaring that it is an excellent idea to purchase a new sofa. To say that the action-awareness of my believing p is non-observational is *not* to say that I cannot have observational awareness of my mental action via introspective or sensory perception, but it is to emphasize that I am not *a mere spectator* of what I am doing mentally.[12] There is a distinctively practical route from which I can gain such awareness of my mental action, which is not available to others. To find out that I form a belief that it is an excellent idea to purchase a new sofa, is *not* a journey

[12] As such, my view is compatible with higher-order theories of consciousness, according to which, very roughly, to be aware of a mental state, is to have a higher-order perception (HOP) or a higher-order thought (HOT) of that state. For defence of a HOP theory of consciousness, see, e.g., Armstrong (1968); Lycan (1996). For defence of a HOT theory of consciousness see, e.g., Rosenthal (1986); Carruthers (2000).

of discovery. I will not be surprised to find out what I am doing mentally, e.g., I am deliberating and coming to believe that it is an excellent idea to purchase a new sofa, as opposed to, say, thinking of a sofa example to make a philosophical point.[13]

So far, I have drawn on the Anscombean idea that S has non-observational knowledge of her believing p if S believes p for reasons, focusing on the phenomenology. I explained what it is for S to be non-observationally aware of her believing p. Now, let us apply the account to the case of believing p in response to the T-practice-based reason. S's believing p is a response to the T-practice-based reason, if S manifests an action-awareness which is distinctively non-observational. S can be aware of her believing p without introspectively perceiving the mental events/states involved. S is not *a mere spectator* of what she is doing mentally. There is a distinctively practical route from which S can gain awareness of her believing p for the T-practice-based reason, which is not available to others. She can do so, simply by making up her mind that p is true!

To further explore the distinctive phenomenology of having non-observational awareness of one's mental action, I will compare William James' description of an episode in which someone recollecting a forgotten name, with Marcel Proust's famous 'madeleine moment'. In both episodes the subject recollects something. But in the former case, the subject engages in a mental *action* of recollecting while in the latter case, a mental *event* of unbidden recollection happens to the subject. The phenomenological contrast between the two cases is particularly striking. So, thinking more carefully about the two cases would help to bring about the distinctive phenomenology of a mental action.[14]

> (RECOLLECTING NAME) 'Suppose we try to recall a forgotten name. The state of our consciousness is peculiar. There is a gap therein; but no mere gap. It is a gap that is intensely active. A sort of wraith of the name is in it, beckoning us in a given direction, making us at moments tingle with the sense of our closeness, and then letting us sink back without the longed-for term ... Everyone

---

[13] This is in agreement with Moran's (2001) influential idea that to account for our first-personal knowledge of what we believe, it is important to pay attention to the role of the subject as an epistemic agent who is capable of taking the stance of rational agency: 'Being the person whose mental life is brought to self-awareness involves a stance of agency beyond that of being a kind of expert witness' (2001:4); '…the stance of the rational agent is the stance where reasons that justify are at issue… Anscombe's question "why" is asking not for what might best explain the movement that constitutes the agent's action, but instead is asking for the reasons he takes to justify his action, what he is aiming at' (2001:127).

[14] This is not, of course, to say that every episode of recollection is either done in a Jamesian or Proustian fashion.

must know the tantalizing effect of the blank rhythm of some forgotten verse, restlessly dancing in one's mind, striving to be filled out with words.' (James 2007: 243–4)

(MADELEINE MOMENT) 'No sooner had the warm liquid mixed with the crumbs touched my palate than a shudder ran through me and I stopped, intent upon the extraordinary thing that was happening to me. An exquisite pleasure had invaded my senses, something isolated, detached, with no suggestion of its origin … Whence did it come? What did it mean? How could I seize and apprehend it? … And suddenly the memory revealed itself. The taste was that of the little piece of madeleine which on Sunday mornings at Combray (because on those mornings I did not go out before mass), when I went to say good morning to her in her bedroom, my aunt Léonie used to give me, dipping it first in her own cup of tea or tisane … As soon as I had recognised the taste of the piece of madeleine soaked in her decoction of lime-blossom which my aunt used to give me… immediately the old grey house upon the street, where her room was, rose up like a stage set…the water-lilies on the Vivonne and the good folk of the village and their little dwellings and the parish church and the whole of Combray and its surroundings, taking shape and solidity, sprang into being, town and gardens alike, from my cup of tea.' (Proust 1996:51-55)

Suppose that I try to recall which village in the Cairngorms I visited last winter and I recall that the name of the village I visited is Braemar. According to James, my awareness of my mental action, i.e., recalling that the name of the village I visited last winter has a distinctive phenomenology: it feels like there is a gap, 'a sort of wraith of the name is in it', beckoning me in a given *direction*. My action-awareness is as of something that *is directed* at the aim for which I engage in that action, that is, to work out which village in the Cairngorms I visited last winter. In this episode of active recollecting, I focus my attention perhaps on visualizing a map of the area, or on a replay of the conversation I had with my friends on the day of travel during which the name of the village was frequently mentioned. I dismiss various other thoughts irrelevant to settling the question of which village I visited last winter. These constraints on my mental activities are self-imposed – I am doing certain things mentally but not others because I take those things to be conducive to the completion of the mental action in question, which is to recollect the name of the village I visited last winter.

By contrast, the episode of unbidden recollection described by Proust has a quite different phenomenology. In that episode, 'I' also come to believe things about 'my' past life in Combray through recollection. However, unlike the previous case, 'my' awareness of the unfolding mental events is not as of something that has a certain direction – it seems spontaneous and open-ended. The unfolding of the series of mental events, from tasting the piece of madeleine soaked in a cup of tea, to visualizing Combray's streets and houses, does not seem to be directed at something. 'I' did not know what to expect or what would be revealed next. Images of Combray appeared suddenly and simply 'sprang into being'. 'My' experience of having introspective awareness of those mental events appears to be as of something that is spontaneous and aimless triggered by the taste of madeleine.

By contrasting with Proust's famous 'madeleine moment', I think the non-observational awareness of one's mental action has two distinctive phenomenological features:

(DIRECTION) S's non-observational awareness of her φ-ing is as of something that has a certain direction – it is moving towards the aim for which she φ-s;

(CONSTRAINT) S's non-observational awareness of her φ-ing is as of something that is subject to some *self-imposed* constraints in the service of achieving that aim for which she φ-s.

Now, applying DIRECTION to the case of believing something for the T-practice-based reason that p is true, my suggestion is that it is part of the phenomenal character of one's experience that it is as of a mental event whose unfolding has a certain direction – it is moving towards the aim of believing something that is true.

Applying CONSTRAINT to the case of believing something for the T-practice-based reason that p is true, we arrive at the following claim: it is part of the phenomenal character of one's experience that it is as of something unfolding is subject to some *self-imposed* constraints in the service of achieving the aim of that action. It feels like as if there are some constraints in place by me so that the unfolding of the events involved smoothly move towards achieving the aim of that mental action, i.e., true belief.

CONSTRAINT is closely related to DIRECTION. Believing for the T-practice-based reason is subject to constraints I impose on my own mental activities. A self-imposed constraint is in contrast with a constraint imposed by external factors. The fact that there is a travel restriction is an external constraint on my bodily movement. The fact that I suffer from an intermittent explosive disorder is an external constraint on how I think about certain

things. Self-imposed constraints, by contrast, are constraints that I impose on my activities regardless of external factors.

Consider a case of revising a belief in critical reflection. Critical reflection, as I am using the notion here, refers to the process of deliberating about p, when one already believes that p. As Adler (2002:286) observes, 'the normal workings of belief is to "blind" us to what might be described from the outside as clues to the contrary.' To believe p is to regard the matter of p as settled. What belief does is to settle the matter and it 'blinds' us from evidence suggesting the contrary. So how is it possible that I re-consider p in self-critical reflection when I already believe that p? A natural answer, it seems to me, is that I do so by bracketing the belief that p, that is, I impose certain constraints on my thinking regarding p. For example, I refrain from using p as a premise in my deliberation about whether p. And I do not dismiss evidence for not-p. These constraints are self-imposed – I am refraining from doing certain things mentally in order to achieve the aim for which I engage in critical reflection, i.e., maintaining true beliefs and removing false beliefs.

But one might have some doubts about the phenomenology I just described. Perhaps in a few cases such as RECOLLECTING NAME, one's mental action manifests the sort of distinctive phenomenology I just sketched. But the phenomenology seems lacking in most cases. One might insist that in forming a belief p in response to the reason that p is true, she cannot tell anything distinctive about the phenomenology of forming the belief p. Her experience does not strike her as of something that has a certain direction – moving towards the aim of believing something true or subject to any self-imposed constraints in the service of achieving that aim.

In response, there are a few plausible explanations as to why one might fail to identify the phenomenology. For instance, in most cases, it is expected that the mental act of believing p in response to the T-practice-based reason happens very quickly. Naturally, one might fail to pay attention to and identify the phenomenology of having an action-awareness that is typically instantaneous or lasts for only a brief period. It is only by reflecting about cases such as RECOLLECTING NAME, which involve actions that last for an extended period, can we bring about the phenomenology of such mental action into a clearer light. But this does not imply that the phenomenology is lacking in cases where the mental action in question is brief.

Moreover, we should distinguish between a lack of a higher-order awareness of one's action-awareness and a lack of action-awareness. The fact that one is not aware of the non-observational action-awareness of her actions does not imply that one lacks non-

observational awareness of her actions. Suppose that I just typed the word 'justice'. Reflecting upon my experience, I was not attentive to my *experience* of typing at that time. But this does not imply that at the time of typing I lack action-awareness of typing the word 'justice'.

Lastly, it is worth emphasizing that S's believing p in response to the T-practice-based reason needs not always be manifested in the phenomenology on my account. The manifestation of the phenomenology is a sufficient but not a necessary condition for reason-responsiveness. Despite a lack of phenomenology, S's believing p can still be a response to the T-practice-based reason if S's belief-formation manifests the relevant reason-responsive dispositions.

Another worry concerns cases where one manifests the described phenomenology of believing for the T-practice-based reason but the T-practice-based reason to believe p does not obtain. Suppose that you come to believe that it is rainy outside when you see that the pavement is wet. Your believing that it is rainy outside manifests a non-observational action-awareness – you can be aware of your belief simply by making up your mind that it is rainy outside when you see that the pavement is wet. Now, suppose that it is not rainy – the T-practice-based reason does not in fact obtain. One might think that in this case, you did not respond to the T-practice-based reason. Yet, you manifested the phenomenology, which according to the account I propose here is a sufficient condition for reason-responsiveness. So, a counterexample!

This, however, is too quick. I agree that the T-practice-based reason to believe that it is rainy does not obtain since it is not rainy. But in this case, the reason you responded to is the fact that the pavement is wet, which does obtain. The phenomenology manifested is the phenomenology of believing that it is rainy in response to the evidential reason that the pavement is wet. While my focus has been developing a hybrid account of reason-responsiveness in the specific case of responding to the T-practice-based reason, the basic idea may very well be generalized. Of course, to fully develop this line of response, I must say more about how to apply the hybrid dispositionalist/phenomenological account to evidential reasons. But it suffices to note here that there is a plausible response to the above case by generalizing the hybrid dispositionalist/phenomenological account to evidential reasons.[15]

---

[15] The hybrid account may also be of value to philosophers who take a reason-responsive approach to rationality (e.g., Kiesewetter 2017; Lord 2018), for there may be a parallel Novice Problem for a

To summarize, in this section I put forward a novel hybrid account of believing something in response to the T-practice-based reason. On this view, S responds to the T-practice-based reason in forming a belief p just in case S manifests the relevant reason-responsive dispositions or manifests an action-awareness of believing p which is distinctively non-observational. The non-observational action-awareness has two features: (1) it is as of something that has a certain direction – moving toward the aim of believing something true and (2) it is as of something that is subject to some self-imposed constraints in the service of achieving that aim. In the next section I will explain how this account can solve the Novice Problem and respond to some objections.

5.4 Solving the Novice Problem

Recall that the Novice Problem arises because according to REG and DRR, a novice of the T-practice cannot be guided by the truth norm of belief, since a novice is someone unable to manifest the relevant dispositions to respond to the T-practice-based reason. This is an unpalatable result for the normativists since a novice of the T-practice is most in need of guidance! I have put the Novice Problem in terms of an incompatible triad:

(REG & DRR)	For any S, p, S's forming a belief p is guided by the truth norm just in case S manifests the relevant dispositions to respond to the T-practice-based reason when the T-practice-based reason is available to S.
(NOVICE)	A novice of the T-practice cannot manifest the relevant dispositions to respond to the T-practice-based reason.
(GUIDANCE)	A novice of the T-practice can be guided by the truth norm of belief.

My diagnosis of the problem was that the manifestation of the relevant reason-responsive dispositions is not a necessary condition for one to respond to the T-practice-based reason, though it might be a sufficient condition. To solve the Novice Problem, I proposed a hybrid dispositionalist/phenomenological account, which I developed and defended in the previous sections. Now, combining REG with HRR, we have:

---

reason-responsive account of rationality understood on purely dispositionalist terms. But I have to leave these issues for future discussion.

(REG & HRR) For any S, p, S's forming a belief p is guided by the truth norm just in case S's believing p either (1) manifests the relevant dispositions to respond to the T-practice-based reason, or (2) manifests a distinctive phenomenology of believing p in response to the T-practice-based reason as a mental action when the reason is available to S.

It is important to emphasize that replacing DRR with HRR still allows REG to meet the four criteria for a viable model of epistemic guidance. First, according to this account, to be guided by the truth norm is for S to form a belief in response to the T-practice-based reason which is to either manifest S's reason-responsive dispositions or the distinctive phenomenology of a reason-responsive mental action. Merely having a belief p caused by the reason that p is true, without manifesting the relevant dispositions or the phenomenology, does not count as being guided by the truth norm.

Second, on REG, guidance is explained in terms of reason-responsiveness. The capacity of reason-responsiveness (either manifesting in the relevant reason-responsive dispositions or in the phenomenology of a reason-responsive mental action) is not understood in terms of guidance. So, REG does not run into a vicious regress.

Third, REG understood in hybrid terms offers a natural explanation of the epistemic worth of beliefs guided by the truth norm. They have epistemic worth because S manifests her capacity to respond to the T-practice-based reason, either by manifesting the reason-responsive dispositions or the phenomenology of a reason-responsive mental action. My account avoids the problem of unworthy guidance.

But most importantly, can REG understood in hybrid terms enable us to explain how the truth norm can guide belief-formation in a way compatible with both Cartesian and Spinozan account of our system of belief-formation? One might worry that my appeal to the phenomenology of the mental action of believing something for truth is in tension with the Spinozan system of belief-formation. Recall that, on the Spinozan account, one forms beliefs automatically when representational contents are made available through cognitive mechanisms such as perception. They are not products of a norm-invoking, effortful inference at person-level. So, when the underlying belief-forming process is Spinozan, how can one's belief-formation have the sort of phenomenology I described above, i.e., that one can be non-observationally aware of believing p by making up one's mind that p is true?

There are a couple of things to say in response. First, one should not read too much into the notion of 'making up one's mind' that I used to explain what is involved in the distinctive non-observational action-awareness. To make up one's mind that p is true is simply to arrive at the belief that p is true without introspecting one's mental states/events. So, making up one's mind in this minimal sense needs not involve deliberative reasoning or making an inference at person-level. For example, I can make up my mind that there is a red cup in front of me simply by looking at the cup. Thus, it is possible to have non-observational awareness of one's belief-formation in both a controlled, reflective process where one first entertains a proposition p and then either arrives at the belief p or rejects it, as well as in an automatic, reflexive process where beliefs are added to one's stock of beliefs without undergoing an evaluation. Second, it is important to emphasize again that on HRR, the phenomenological condition is a sufficient but not a necessary condition. Even if one's belief-formation lacks the relevant phenomenology, it may still be guided by the truth norm if it manifests the relevant reason-responsive dispositions.

With these clarifications in place, we can see that the Novice Problem does not arise because REG understood in hybrid terms is compatible with both NOVICE and GUIDANCE. According to REG and HRR, a novice of the T-practice is guided by the truth norm if she can be aware of her believing p via a non-observational, practical route, by making up her mind that p is true. Her non-observational action-awareness is as of something that has a certain direction – moving toward the aim of believing something true and is subject to some self-imposed constraints in the service of achieving that aim. So, a novice can be guided by the truth norm of belief on REG understood in hybrid terms.

In short, REG with HRR is an attractive and viable account of how the truth norm can guide our belief-formation and it enables us to provide a satisfactory response to the guidance challenge. A novice of the T-practice can form a belief in response to the T-practice-based reason and be guided by the truth norm of belief.

5.5 Conclusion

In Chapter 5, I discussed the Novice Problem. The problem arises for it seems that a novice of the T-practice cannot be guided by the truth norm of belief on REG understood in dispositionalist terms. I provided a solution to the Novice Problem and in doing so I developed a novel hybrid dispositionalist/phenomenological account of what it takes to respond to the T-practice-based reason. My central idea is that believing p in response to the T-practice-based reason is a mental action. For one to respond to the T-practice-based is for

one to either manifest the relevant dispositions to respond to the T-practice-based reason or manifest a distinctive phenomenology of the mental action of believing something for the T-practice-based reason. I showed that the hybrid account can deal with the Novice Problem. REG understood in hybrid terms offers an attractive response to the guidance challenge.

Part III. Can the Truth Norm Make Any Claim Regarding What to Believe?

Summary of Part II and Preview of Part III

In Part II, I addressed the guidance challenge. Norms must be capable of providing guidance. On the reason-based framework, if the truth norm of belief can provide guidance, then it must be possible for us to respond to the T-practice-based reason. In Chapter 4, I argued that the existing response to the No Guidance argument is inadequate considering the emerging empirical evidence about our system of belief-formation. I motivated an alternative response to No Guidance, and presented, in rough outline, a reason-responsive account of epistemic guidance (REG). In Chapter 5, I defended the view that the truth norm of belief can provide guidance on REG and offered a solution to the Novice Problem. I suggested that for one to respond to the T-practice-based (and thus be guided by the truth norm) is for one either to manifest the relevant dispositions to respond to the T-practice-based reason or to manifest a distinctive phenomenology of the mental action of believing something for the T-practice-based reason.

In Part III, I will turn to the weighting challenge. Norms are the sort of things that tell us what we ought to or may do. What sort of claim, then, does the truth norm of belief make on us? In Chapter 6, I will consider and criticize two existing views: the obligatory and the permissible account of the truth norm. Both views are what I will call *invariantist* views, for they claim that the T-practice-based reason has the same requiring/permitting weight across all contexts. Both views face counterexamples and thus fail to meet the weighting challenge. In Chapter 7, I will defend a *variantist* account of the truth norm, according to which the weight of the T-practice-based reason can vary from context to context. I will argue that variantism offers an attractive response to the weighting challenge compared to other existing alternatives.

6. The Weighting Challenge

Norms are the sort of things that can make a claim on us regarding what to do: they tell us what one ought to or may do. So, if the truth norm of belief is a genuine norm, it must be capable of making a claim on us regarding what we ought to or may believe. What claim then, does the truth norm of belief make on us?

According to one view, the truth norm of belief is obligatory: for any S, p, S *ought to* believe p if and only if p is true (e.g., Gibbard 2003, 2005; Wedgwood 2002, 2007, 2013a, 2013c; Shah 2003, Shah and Velleman 2005). This once orthodox view has been heavily criticized in recent discussion on three grounds: (1) it makes excessive demands on us regarding what we ought to believe; (2) it makes demands on us regarding true blindspot propositions that cannot be satisfied; and (3) it makes demands on us that may clash with our epistemic obligation to believe in accordance with evidence. These problems have motivated the permissible account of the truth norm, according to which for any S, p, S *may* believe p if and only if p is true (e.g., Whiting 2010, 2012, 2013).

In this chapter, I will argue that both views are mistaken and motivate an alternative, variantist account, according to which the sort of claim the truth norm of belief can make on us varies from context to context. I will begin by discussing three problems facing the obligatory account of the truth norm, which have been thought to favour the permissible account of the truth norm (Sect. 6.1). I will then argue that the permissible account of the truth norm does not fare any better than the obligatory account. I offer a diagnosis of the three problems on a reason-based normative framework. They are what I will call *invariantist* views, for they claim that the T-practice-based reason has the same requiring/permitting weight across all contexts. Both invariantist views face three kinds of counterexamples (Sect. 6.2). Next, I will consider two natural moves invariantists may make in response and argue that there is no quick fix of their problems (Sect. 6.3 & Sect. 6.4). Thus, the invariantist views fail to meet the weighting challenge. This motivates an alternative, variantist account of the truth norm, which will be developed and defended in the next chapter.

6.1 The State of the Debate

There are two prominent families of views in the current literature regarding the weighting challenge. According to the obligatory account of the truth norm (e.g., Gibbard 2003, 2005; Wedgwood 2002, 2007, 2013a, 2013c; Shah 2003, Shah and Velleman 2005):

($\mathcal{T}_O$) For any S, p, S ought to believe p if and only if p is true.

Versions of $\mathcal{T}_O$ have once been the orthodox view, for many philosophers take truth to be the standard of correctness for belief and understand 'correctness' as a normative notion concerning what one ought to do. Since a belief is correct if and only if it is true, intuitively, one might think one thereby ought to believe p if and only if p is true.[1]

The obligatory account is in contrast with a weaker, permissible norm of truth, which is thought to be more plausible (e.g., Whiting 2010, 2012, 2013) :

($\mathcal{T}_P$) For any S, p, S *may* believe p if and only if p is true.

$\mathcal{T}_P$ seems just as normatively interesting as $\mathcal{T}_O$, for it not only tells us what we are permitted to believe, it also tells us what we ought *not* to believe (Whiting 2010: 216-217). $\mathcal{T}_P$ can play a significant role in our normative theorizing just like $\mathcal{T}_O$.

In recent discussion, $\mathcal{T}_O$ has been widely criticized on three grounds: (1) $\mathcal{T}_O$ seems to make excessive demands on us regarding what we ought to believe; (2) $\mathcal{T}_O$ seems to make unsatisfiable demands on us regarding true blindspot propositions; and (3) $\mathcal{T}_O$ seems to make demands on us that may clash with our epistemic obligation to believe in accordance with evidence. These problems are thought to motivate $\mathcal{T}_P$. Let us go through them in turn.

6.1.1 The Problem of Excessive Demands

One common objection to the obligatory account is that it makes excessive demands on us regarding what we ought to believe (e.g., Bykvist & Hattiangadi 2007; Glüer & Wikforss 2013). The claim that for any p, the truth of p is a requiring reason for us to believe p leads to an 'explosion' of epistemic obligations which are excessive given that we are ordinary epistemic agents with finite cognitive powers. Since there are infinitely many truths in the world (and some of which are far too complex for most humans to believe), we cannot, surely, believe every single one of

[1] Not all philosophers agree that 'correctness' is a normative notion concerning what one is *required* to do. Whiting (2010), for instance, has argued that 'correct' should be understood as a normative notion concerning what one is *permitted* to do. To begin with, it seems possible to judge that φ-ing is correct without judging that one is required to φ but it seems impossible to judge that φ-ing is correct without judging that one is permitted to φ. Furthermore, according to Whiting, our linguistic data suggest that 'correctness' serves mainly as a license. For example, in saying that 'Nf3 is a correct move', I do seem to express the claim that one is required to move the knight to f3, but rather, I simply mean that Nf3 is a permissible move.

them. As such, the obligatory account of the truth norm clashes with the intuitive principle 'ought' implies 'can' (OIC).[2] In other words, the following three claims are jointly incompatible:

(OIC)	For any S, φ, necessarily, if S ought to φ then S can φ.
(Limited Capacity)	For some S, p, S cannot believe p when p is true.
($\mathcal{T}_O$)	For any S, p, S ought to believe p if and only if p is true.

According to the obligatory account of the truth norm, for any S and any proposition p, S has an obligation to believe p when p is true. It follows then, by OIC, for any S, p, if p is true, S can believe p. However, Limited Capacity says that there are cases where when p is true, S cannot believe p. So, one of the three claims has to go. Call this the problem of Excessive Demands.[3]

The permissible account of the truth norm seems to avoid the problem of Excessive Demands all together, after all, it says that for any S, p, the truth of p *permits* S to believe p but falls short of *requiring* it. There is no 'explosion' of epistemic obligations. It is not the case that we are required to believe every truth that is out there. The permissible account of the truth norm is compatible with OIC and Limited Capacity. Since both Limited Capacity and OIC are plausible,

---

[2] OIC has been a historical influential and widely employed principle. For example, historically, it plays a significant role in Kant's argument for the immortality of the soul. Very roughly, according to Kant (1997), complete conformity with the moral law is only possible in an endless progress. Since we ought to completely conform with the moral law, endless progress is possible and hence the endlessly continuation of the existence of personality of the same rational being, i.e., immortality of the soul (5:122). It is used to derive the Principle of Alternative Possibilities, according to which one is morally blameworthy for φ-ing only if one could have done something other than φ (e.g., Copp 1997; Widerker 1991). It is also used to show that it is impossible that a person has two moral obligations but cannot satisfy both. Hence there are no genuine moral dilemmas (e.g., Gowans 1987; Mason 1996). It is employed to refute probabilistic theories of rationality. On any probabilistic theory of rationality, probabilistic coherence entails that one is certain of all necessary truths. But since one cannot be certain of all necessary truths and given OIC, probabilistic theories of rationality are false (e.g., Smithies 2015).

[3] Wedgwood (2013a) responds to this problem by proposing a restricted version of the obligatory account of the truth norm: ($\mathcal{T}_O$*) For any S, p, S ought to believe p if and only if p is true *and S considers p* (See also, Feldman 2000, Greenberg 2020). There is no 'explosion' of epistemic obligations since for those true propositions that are not under consideration, we are not obliged to form a belief about them. Furthermore, it may be argued that as long as I can consider a proposition, I can form a belief about that proposition, for no matter how complex it is, I can form a belief ostensibly, e.g., I can believe *that thing* (thanks to Alessandra Tanesini for raising this point in conversation). While I have my doubts about whether it is possible for ordinary epistemic agents to believe every true proposition we consider, a more urgent problem with Wedgwood's restricted version is that it under-generates epistemic obligations. Consider epistemic wrongdoings such as self-deception. Some truths can hurt. One common strategy to deceive oneself about undesirable, unpleasant truths is to avoid considering them and pretend some desirable, pleasant falsehoods to be true (e.g., Wei 2020). But it seems wrong to say that in these cases, the T-practice-based reason does not have a requiring weight and the self-deceptive agent is not obliged to believe those truths because they are not under consideration. For this reason, I do not think restricting epistemic obligations to doxastic attitudes one considers is the right response to the problem of Excessive Demands.

the fact that they are incompatible with $\mathcal{T}_{\mathrm{O}}$ but compatible with $\mathcal{T}_{\mathrm{P}}$ seems to be a good reason to reject the former and adopt the latter.

6.1.2 The Problem of Blindspot Propositions

The obligatory account of the truth norm has also been objected on the ground that it requires us to believe true blindspot propositions, but such demands cannot be satisfied (e.g., Bykvist & Hattiangadi 2007, 2013). Schematically, a blindspot proposition α is of the following form:

> (Blindspot) For any S, α: necessarily, if α is true, then S does not believe α and necessarily, if S believes α, then α is false.[4]

Necessarily, α's being true is incompatible with S's believing α. Here is an example. Consider the complex proposition 'it is rainy and nobody believes that it is rainy'. If that complex proposition is true, then necessarily, both of its conjuncts are true. Since anyone who believes 'it is rainy' would make the proposition 'nobody believe that it is rainy' false, so necessarily, nobody believes 'it is rainy and nobody believes that it is rainy'. And for any S, if S believes 'it is rainy and nobody believes that it is rainy', then necessarily, she believes both conjuncts are true. Since she believes the first conjunct is true, then necessarily, the second conjunct is false, so the proposition 'it is rainy and nobody believes that it is rainy' is false. A proposition like 'it is rainy and nobody believes that it is rainy' is a blindspot proposition.

Why do blindspot propositions pose a problem for the obligatory account of the truth norm? Suppose that the complex proposition 'it is rainy and nobody believes that it is rainy' is true. According to $\mathcal{T}_{\mathrm{O}}$, S is required to believe 'it is rainy and nobody believes that it is rainy'. But if S believes this proposition, then S believes both conjuncts, including the first conjunct 'it is rainy'. Since S believes that 'it is rainy', the second conjunct is false. So, the complex proposition 'it is rainy and nobody believes that it is rainy' is false. So, S is required by the same norm not to believe that proposition! It seems then, with respect to a blindspot proposition α, it is impossible for one to satisfy one's epistemic obligation to believe α when α is true, while its being true that one has an obligation to believe α. Bykvist and Hattiangadi argue that $\mathcal{T}_{\mathrm{O}}$ is false because it cannot be satisfied, and hence, according to them, cannot exist:

---

[4] According to Sorensen (1988), blindspot propositions are not restricted to belief but can be applied to any given propositional attitude. But for our purpose, we will focus on blindspot propositions for belief.

> 'The problem here is not that the proposition cannot be believed, but that the obligation to believe p cannot be satisfied. So this is not a violation of the principle that 'ought' implies 'can'. Rather, it is a violation of the principle that 'ought' implies 'can satisfy', which says that if you ought to believe that p, then it is logically possible for you to discharge or satisfy this ought. Or, more generally: 'Ought' implies 'can satisfy'. If you ought to A, then it is logically possible for you to A while its being true that you ought to A. Now, the principle that 'ought' implies 'can satisfy' seems as plausible as the principle that 'ought' implies 'can'. Just as one cannot have an obligation to do what is impossible to do, one cannot have an obligation that it is impossible to satisfy.' (Bykvist and Hattiangadi 2013:109)

Since the obligatory truth norm cannot be satisfied with respect to true blindspot propositions, the obligatory account of the truth norm is jointly incompatible with the principle of 'ought' implies 'can satisfy' (OICS) and the fact that there are true blindspot propositions. Call this the problem of Blindspot Propositions.

By contrast, the permissible account of the truth norm seems to avoid the problem of Blindspot Propositions. After all, the permissible account of the truth norm says that for any S, p, the truth of p *permits* S to believe p but falls short of *requiring* it. There is no epistemic obligation to believe true blindspot proposition: S is permitted to not to believe α. The fact that $\mathcal{T}_{\mathrm{O}}$ but not $\mathcal{T}_{\mathrm{P}}$ violates OICS seems to be a good reason to reject the former and adopt the latter view (e.g., Whiting 2010:218-220).

6.1.3 The Problem of Conflicting Norms

Another problem for the obligatory account of the truth norm is that in some cases, it makes demands on us that clash with our epistemic obligation to believe in accordance with evidence (e.g., Hattiangadi 2010; Glüer & Wikforss 2013). To fix ideas, consider the following norm of evidence:

> (No Evidence) For any S, p, S ought to refrain from believing p if S has no evidence supporting p.[5]

[5] No Evidence might be objected on the ground that what one ought to or may believe is (partly) determined by non-evidential, pragmatic considerations (e.g., Rinard 2015, 2018, 2019; Reisner 2018). For example, if you want to accomplish an important task, then even if you have no evidence suggesting that you will accomplish that task, you are permitted to believe you will succeed anyway (e.g., Marušić 2015). I

Since the obligatory account of the truth norm implies that for any p, S, S ought to believe p when p is true, it clashes with No Evidence in cases where p is true, but S has no evidence supporting p. In such cases, S is required to believe p by $\mathcal{T}_O$ but S is also required to refrain from believing p by No Evidence. To illustrate, suppose that I toss a fair coin. The coin has landed heads. But I have not revealed the coin to you. According to the truth norm, you ought to believe that the coin has landed heads. According to No Evidence, you ought to refrain from believing that the coin has landed heads for you lack evidence. In this case, it seems that you are both required to believe that the coin has landed heads and refrain from believing that the coin has landed heads. But this is not something you can do: you cannot both believe p and refrain from believing p. The obligatory account of the truth norm is jointly incompatible with OIC and No Evidence. Call this the problem of Conflicting Norms.[6]

By contrast, the permissible account of the truth norm seems to avoid the problem of Conflicting Norms. Since the permissible account says that for any S, p, the truth of p *permits* S to believe p but falls short of requiring it, it is not the case that S is required to believe p when p is true but unsupported by S's evidence. You are not required to believe that the coin has landed heads by $\mathcal{T}_p$. There is no clash between $\mathcal{T}_p$ and No Evidence. Since both No Evidence and OIC are intuitively plausible principles, the fact that they are incompatible with $\mathcal{T}_O$ but compatible with $\mathcal{T}_p$ seems to be a good reason to reject the former and adopt the latter.

6.2 Against the Permissible Norm of Truth

The three arguments we just considered, taken together, seem to make a persuasive case against $\mathcal{T}_O$ and in favour of $\mathcal{T}_p$. However, I think upon a closer look, $\mathcal{T}$p is not preferable to $\mathcal{T}_O$. In this section, I will first show that $\mathcal{T}_p$ in fact faces analogous problems as $\mathcal{T}_O$. I will then offer a diagnosis of the three problems on the reason-based normative framework we are working with.[7]

---

will leave this debate aside. To get the argument – that since $\mathcal{T}_O$ but not $\mathcal{T}_p$ clashes with No Evidence, $\mathcal{T}_p$ is preferable to $\mathcal{T}_O$ – going, it suffices to say that No Evidence is *prima facie* plausible.

[6] One response to this problem is to endorse perspectivism, according to which what one ought to believe is subject to one's epistemic perspective. We will return to this view in Chapter 7, Sect. 7.4.2.

[7] This section contains revised material from a previously published article in Wei (2019).

6.2.1 The Problem of Excessive Demands Revisited

It is commonly accepted that 'ought' and 'may' are dual concepts and 'may' is commonly defined as follows: may x $=_{def}$ ¬ ought ¬x.[8] It is not difficult to see that $\mathcal{T}_P$ implies an obligatory norm of falsity: for any S, p, S has an epistemic obligation to refrain from believing p when p is false. $\mathcal{T}_P$ can be broken into two conditionals:

($\vec{\mathcal{T}_P}$) For any S, p, if p is true, then S may believe p.

($\overleftarrow{\mathcal{T}_P}$) For any S, p, S may believe p only if p is true.

Using contraposition (and assuming bivalence, i.e., that there are only two truth values), $\overleftarrow{\mathcal{T}_P}$ is equivalent to:

For any S, p, if p is false, then it is not the case that S may believe p.

Hence, given that may x $=_{def}$ ¬ ought ¬x, $\overleftarrow{\mathcal{T}_P}$ is equivalent to the following claim:

($\mathcal{F}_O$) For any S, p, S ought to refrain from believing p if p is false.

Thus, $\mathcal{T}_P$ entails an obligatory falsity norm of belief such that, for any S, p, if p is false, S ought to refrain from believing p. $\mathcal{F}_O$ faces an analogous problem of excessive demands, since plausibly there are some falsehoods, such that it is not possible for one to refrain from believing them. $\mathcal{T}_P$ is therefore jointly incompatible with the following two claims:

(OIC)	For any S, φ, necessarily, if S ought to φ then S can φ.
(Limited Capacity*)	For some S, p, S cannot refrain from believing p when p is false.

The permissible norm of truth implies that for any S and any false proposition p, S has an *obligation* to refrain from believing p. By OIC, for any proposition p, if p is false, then S can refrain from believing p. However, Limited Capacity* says that there are cases in which S cannot refrain from believing p when p is false. So, at least one of the three claims has to go.

Whiting (2013:125-126) rejects Limited Capacity*. He invites us to consider a case where someone is said to be psychologically unable to refrain from believing that there are aliens. Suppose that there are no aliens. Does an example like this show that Limited Capacity* is true? Whiting

[8] Here I follow the notations in the standard deontic logic (e.g., McNamara 2010).

thinks not. First, he complains that the relevant modality of 'can' involved in OIC is weaker than psychological possibility. He suggests that 'ought to φ' implies that it is '*humanly possible* to φ'. Since it is humanly possible to for one to refrain from believing that there are aliens, the person in the example *can* refrain from believing that there are aliens even if he is psychologically unable to do so. Second, Whiting argues that no critic of $\mathcal{T}_p$ has shown that there are cases where if p is false, it is *humanly impossible* to refrain from believing that p. Finally, he claims that even if the critic of the permissible account of the truth norm can show that there are such cases, there is a further question whether the attitude S has towards p counts as a genuine belief.

The question regarding the modality of 'can' is indeed an important one. However, how should we understand the notion of 'humanly possible' to φ? And can we make sense of something that is humanly possible but not psychologically possible?

According to the standard interpretation of 'can φ', one can φ just in case one (1) has the ability to φ and (2) has the opportunity to exercise that ability to φ.[9] On one influential view, one has an opportunity to φ if there is a non-zero objective chance to φ assigned by the relevant psychological laws, where psychological laws are laws that are broadly based on folk-psychology and deal with agent's actions and attitudes.[10] On this common interpretation of 'can φ' as having the ability and opportunity to φ, given the psychological laws governing agent's actions and attitudes, we have at least some reasonable grasp of what 'can φ' amounts to, broadly based on folk-psychology. By contrast, Whiting does not explain his notion of 'humanly possible' to φ. On the face of it, whether it is 'humanly possible' to φ would depend on the kind of creature we are, empirically speaking. And we are creatures subject to psychological laws. If that is right, a natural way to flesh out what is 'humanly possible' to φ is just the standard interpretation of *can φ*. It is *humanly possible* for S to φ just in case S has the ability and opportunity to φ, given the psychological laws governing our actions and attitudes.

Having said that, I agree with Whiting that critics of the permissible account of the truth norm are yet to show that Limited Capacity* is true. The case Whiting offers on behalf of his critics – that of a person who cannot refrain from believing that there are aliens does not lend much support to Limited Capacity*. It is hardly convincing that, as far as how the case is described,

[9] Such formulation is widely adopted in the debate concerning OIC. See, e.g., Copp (2008); Granham (2011); Mizrahi (2009, 2012); Vranas (2007); Wedgwood (2013b).
[10] See, Wedgwood (2013b), we shall return to this in Sect.6.4.

that the person genuinely cannot refrain from believing that there are aliens. I now turn to the task of offering three more persuasive cases in support of Limited Capacity*.

First, some beliefs might be deeply integrated in our psychological make-up that we cannot refrain from having them. Consider forms of clinical delusions, e.g. patients with Capgras delusion cannot refrain from believing that a close relative has been replaced by an impostor, often due to cognitive failure including abnormal perceptual experiences (as a result of a malfunctioning face recognition system) and possibly also with a deficit in their belief evaluation system.[11] Now, of course, few of us suffer from clinical delusions, yet I think some of our core beliefs may be psychologically impossible to shake off in a rather similar way as a result of how we are hard-wired to perceive the world. In fact, many philosophical theories, if correct, would render some of our core beliefs false. For instance, if error theories about mathematics and ethics are correct, none of our mathematical and ethical beliefs are literally true.[12] If the B-theory of time is correct, then the passage of time is an illusion and the present is not ontologically privileged.[13] And yet, arguably, we cannot refrain from having beliefs about temporal experiences, that 2+2=4, or that murder is wrong.

Second, it is not always within our power to avoid falsity if all evidence available to our epistemic community supports the false belief in question. For example, we might say that the best evidence available to the ancient Greek supports the claim that Phosphorus and Hesperus are two different celestial bodies. Given the restricted epistemic circumstances back then, arguably one cannot revise the false belief that Phosphorus and Hesperus are two distinct celestial bodies. Similarly, some of our current scientific beliefs may turn out to be false, yet we may not be able to revise them now if they are supported by our best evidence. Of course, given the development of science and technology, more evidence will become available and we will be able to spot more falsehoods and revise our beliefs accordingly. Indeed, this is the story of our scientific progress. However, for any given time, our epistemic position is always limited, and we cannot revise our false beliefs if they are supported by the best evidence available to our community at the time.

Third, some propositions are deeply integrated in our epistemic life, such as the so-called cornerstone propositions (e.g., that there is an external world, that I am not a brain in a vat are cornerstone propositions).[14] Arguably, we cannot refrain from accepting them despite the

---

[11] For a recent overview of neuropsychological accounts of delusions, see, Bortolotti (2009).
[12] See, notably, Field (1989).
[13] For an influential account of B-theory of time, see, for instance, Sider (2001).
[14] The concept of cornerstone proposition is first coined in Wright (2004), which is inspired by Wittgenstein's idea of hinge proposition in *On Certainty* (1969). Roughly, a proposition is a cornerstone just in case, if we had no warrant for that proposition we could not rationally claim warrant for belief in any

possibility that they are false. If I were a brain in a vat, then those cornerstone propositions would be false. Yet, can I genuinely refrain from believing those cornerstone propositions? Perhaps in an epistemology seminar I can momentarily suspend *judgments* about cornerstone propositions while entertaining the sceptical scenarios. However, it is hard to imagine that we can refraining from *believing* cornerstone propositions if we were to live a normal life. It is not possible for me to not to be disposed to act and think as if it were true that there is an external world and that I am not a brain in a vat. If I did not believe that I am not a brain in vat, I would not be able to have the ordinary empirical beliefs which are crucial for me to navigate through the world. Of course, the point here is not to claim that scepticism is true. Rather, the point is to emphasize that there are some propositions at the core of our belief system that we cannot refrain from believing as far as we are engaging in any believing at all and as we carry on our daily life, given the kind of creatures we are. As such, if scepticism were true, we would not be able to refrain from believing false cornerstone propositions.

In short, the underlying thought is this: given the psychological and cognitive constraints, and the fact that the world is not always cooperative, we cannot always refrain from believing things when they are false. Given that we are finite epistemic agents, we are limited not only with respect to what we can believe but also what we can refrain from believing.

Now you might point out that Whiting could still maintain that even if we have shown that Limited Capacity* is true, there is a further question as to whether the attitude in the cases I discussed is in fact a belief. Suppose that I cannot refrain from believing, say, that 2+2=4, even in the presence of overwhelming evidence that mathematical fictionalism is true, then, it may be argued that my attitude towards the proposition 2+2=4 is not that of belief.

I do not see how one can maintain this point without presupposing that it is constitutive of belief that it is subject to a norm of truth along the line we explored in Chapter 2. We have already seen that arguments for such the constitutive account are unpersuasive. But without appealing to this constitutive claim it is hard to see why my attitude towards that 2+2=4 fails to be a belief, as long as the attitude plays the characteristic role belief plays in my mental economy (Sect. 1.2.1).

Thus, the permissible account of the truth norm in fact faces analogous problem of Excessive Demands. $\mathcal{T}_{\mathrm{p}}$ makes excessive demands on us regarding what we ought to refrain from believing.

---

proposition about the empirical world. My use of cornerstone proposition simplifies the detail of Wright's account.

6.2.2 The Problem of Blindspot Propositions Revisited

The permissible account of the truth norm also faces an analogous problem of Blindspot Propositions. Recall that a blindspot proposition has the following structure:

> (Blindspot) For any S, α: necessarily, if α is true, then S does not believe α and necessarily, if S believes α, then α is false.

Suppose that the complex proposition 'it is rainy and nobody believes that it is rainy' is true. $\mathcal{T}_{\mathrm{p}}$ implies that, for any S, S may believe 'it is rainy and nobody believes that it is rainy'. But if S believes that proposition, then necessarily S believes both conjuncts, including the first conjunct 'it is rainy'. If S believes 'it is rainy', then the second conjunct is false. So, if S believes that 'it is rainy and nobody believes that it is rainy', this proposition is false. We have seen that the permissible account of the truth norm also implies that for any S, p, S has an epistemic obligation to refrain from believing p when p is false. So, if the permissible account of the truth norm is true, it says that it is permissible to believe true blindspot propositions. But if I were to believe those propositions, I would believe things that I am not permitted to believe. Bykvist and Hattiangadi explain why this result is unpalatable:

> So what we have here is a permission which, if acted upon, would unavoidably change into a prohibition. And this seems fishy. Here's a vivid illustration of the fishiness: imagine you went to a (fish!) restaurant that offers an all-you-can-eat buffet. You pay, and tuck in, but as you do, the waiters come running and explain that you are permitted to eat as much as you want only if you do not eat as much as you want, whereas if you do eat as much as you want, you are forbidden to do so. (Bykvist and Hattiangadi 2013:113)

Of course, proponents of $\mathcal{T}_{\mathrm{p}}$ could bite the bullet, after all, the fishy consequences are limited to blindspot propositions only. However, dialectically, this objection is significant, for it shows that $\mathcal{T}_{\mathrm{p}}$ does not enjoy an edge of advantage over $\mathcal{T}_{\mathrm{O}}$ when it comes to the problem of Blindspot Propositions.

6.2.3 The Problem of Conflicting Norms Revisited

$\mathcal{T}_{\mathrm{p}}$ also faces an analogous problem of Conflicting Norms. $\mathcal{T}_{\mathrm{p}}$ implies $\mathcal{F}_{\mathrm{O}}$: one is required to refrain from believing p when p is false. But the epistemic obligation to refrain from believing false propositions can come into clash with our epistemic obligation to believe in accordance with

evidence. Consider the following norm of evidence:

(Evidence) For any S, p: S ought to believe p if S's evidence supports p.[15]

Since the $\mathcal{T}_p$ implies that for any p, S, S ought to refrain from believing p when p is false, it clashes with Evidence in cases where S's evidence supports p, but p is false. In such cases, S is required to believe p by Evidence, and S is also required to refrain from believing p by $\mathcal{T}_p$. But S cannot both believe p and refrain from believing p! $\mathcal{T}_p$ is jointly incompatible with OIC and Evidence. So, at least one of the three claims has to go.

But of course, one might find Evidence too strong and opt instead for a weaker norm of evidence:

(Evidence*) For any S, any p: S may believe p if S's evidence supports p.

But Evidence* does not help proponents of $\mathcal{T}_p$, on the contrary, it worsens the problem. For Evidence* contradicts $\mathcal{T}_p$ in cases where S's evidence supports p but p is false. To illustrate, suppose that Fin was hiking in the Cairngorms National Park and saw a large white object that looked like a mountain hare. Fin's evidence suggests that he saw a mountain hare. But in fact, what Fin saw was the local farmer's pet rabbit. In this case, according to $\mathcal{T}_P$, Fin is not permitted to believe that he saw a mountain hare, since it is false. But according to Evidence*, Fin is permitted to believe that he saw a mountain hare since his evidence supports it. A contradiction! So, either $\mathcal{T}_p$ or Evidence* has to go.

Again, dialectically, this objection is significant, for it shows that $\mathcal{T}_p$ is not preferable to $\mathcal{T}_O$ when it comes to the problem of Conflicting Norms and so in that respect $\mathcal{T}_p$ does not fare any better than $\mathcal{T}_O$ which it aims to replace.

6.2.4 The Weight of the T-practice-based Reason: A Diagnosis

On the reason-based normative perspective framework we are working with, the debate between the obligatory account of the truth norm and the permissible account of the truth norm is a debate

---

[15] See Feldman (1988) and Conee & Feldman (2004) for their influential defence of Evidence. Notice that, my point here is to show that $\mathcal{T}_p$ does not fare better than $\mathcal{T}_O$ with respect to the problem of Conflicting Norms, since they both issue demands that can clash with our epistemic obligation to believe in accordance with evidence. To get this argument going, it suffices to say that Evidence is *prima facie* plausible.

about the weight of the T-practice-based reason in making a claim on us regarding what we ought to or may believe.

The idea that reasons have weight is familiar from our everyday talk about reason (see, e.g., Maguire & Lord 2016). When one stands in a relation to a reason, the extent to which a reason counts in favour of φ-ing comes in various degrees: the stronger the normative support it has for one to φ, the weightier the reason is. For example, the fact that you like ice cream is a reason for you to have ice cream for dessert, but the fact that your doctor advices against it is a weightier reason for you not to have ice cream for dessert.

As I have said in Sect. 1.2.3, we can distinguish between reasons that are weighty enough to *require* certain responses and reasons that are weighty enough to *permit* certain responses but falling short of *requiring* them.[16] Intuitively, a reason R is weighty enough to require S to φ, if S ought to φ and is blameworthy for not φ-ing without good excuses when R obtains. By contrast, a reason R is weighty enough to permit (but falls short of requiring) S to φ, if S may φ and S is not blameworthy for not φ-ing when R obtains.

Consider a few examples. Climate emergency is an existential threat to humanity. That seems to permit me to participate in civil disobedience in protest of my government's inaction, but not to require it. Although I may join the protest, it does not seem to be the case that I ought to participate in civil disobedience. I am not blameworthy for not doing so. My uncle is allergic to nuts, that seems to require him to refrain from eating nuts, not merely to permit him to refrain from doing so. He ought to refrain from eating nuts. Of course, his action may be excusable if, for example, he does not know the food he eats contains nuts. But without a good excuse, he is blameworthy for failing to refrain from eating nuts. Alice has been in contact with a Covid-19 patient and she has developed a fever. That seems to permit me to believe that Alice has Covid-19, but not to require it, after all, the evidence also supports the hypothesis that Alice has a flu. Later she tested positive for Covid-19 and the test was accurate. That seems to require me to believe that Alice has Covid-19, not merely to permit it. I ought to believe that Alice has Covid-19.[17]

---

[16] The requiring/permitting distinction is widely drawn among philosophers, though they might use different terms such as 'demanding/justifying', 'decisive/sufficient' (see, e.g., Dancy 2004; Gert 2003, 2007; Scanlon 2014; Kiesewetter 2017:8; Whiting 2021).

[17] As we will see in Chapter 7, the distinction between requiring and permitting weight is more complicated than what the above examples illustrate. But for now, the distinction suffices for us to put the obligatory and permissible account of the truth norm in terms of the weight of the T-practice-based reason.

The obligatory and permissible account of the truth norm, on our framework, are what I will call *invariantist* accounts, for they claim that the T-practice-based reason has a certain weight across all contexts. According to $\mathcal{T}_O$, the T-practice-based reason has a requiring weight regarding what to believe: for any S, p, the T-practice-based reason requires S to believe p when p is true. Our discussion of the problem of Excessive Demands, Blindspot Propositions and Conflicting Norms illustrate three kinds of cases in which the T-practice-based reason fails to be a requiring reason for one to believe p when p is true:

(a) Cases involving propositions which we cannot believe;
(b) Cases involving blindspot propositions;
(c) Cases involving propositions unsupported by one's evidence.

But as we have seen, the alternative invariantist account, i.e., $\mathcal{T}_P$, does not fare any better. According to $\mathcal{T}_P$, the T-practice-based reason has a permitting weight regarding what to believe and a requiring weight regarding what not to believe: for any S, p, the T-practice-based reason *permits* S to believe p and *requires* S to refrain from believing not-p when p is true (i.e., when not-p is false). But there are analogous cases in which the T-practice-based reason either fails to be a permitting reason for one to believe p when p is true or fails to be a requiring reason for one to refrain from believing not-p when p is true:

(a) Cases involving propositions which we cannot refrain from believing;
(b) Cases involving blindspot propositions;
(c) Cases involving propositions unsupported by one's evidence.

The problem of Excessive Demands, Blindspot Propositions and Conflicting Norms arise because there are cases in which the T-practice-based reason does not permit one to believe p or require one to refrain from believing not-p when p is true, since the T-practice-based reason does not have an invariant weight across all situations. Whether the T-practice-based reason has a requiring/permitting weight can vary depending on the circumstances. Call this view the variantist account of the truth norm.

But before spelling out in more detail just how the weight of the T-practice-based reason may vary and how variantism meets the weighting challenge, we must first consider and respond to some potential moves an invariantist might make to resist the arguments we have considered so far.

6.3 Alternative Theories of Doxastic Oughts

One might think that we can make the problems of Excessive Demands, Blindspot Propositions and Conflicting Norms go away by appealing to theories of doxastic oughts unconstrained by OIC and OICS. In this section, I will consider two such proposals. They are put forward by Feldman (2000) and Chrisman (2008) originally as a response to a well-known problem facing epistemic deontology from doxastic involuntarism, but may be of use to the invariantists in the present context.[18] I will argue that neither can be made to work to help the invariantist accounts of the truth norm.

Feldman (2000) suggests that doxastic obligations should be understood as a kind of 'role oughts', which describe what is the right way to play a certain role. According to Feldman, 'role oughts' are based on what counts as good performance appropriate to the role in question:

> 'There are oughts that result from one's playing a certain role or having a certain position. Teachers ought to explain things clearly. Parents ought to take care of their kids. Cyclists ought to move in various ways. Incompetent teachers, incapable parents, and untrained cyclists may be unable to do what they ought to do [...] I suggest that epistemic oughts are of this sort-they describe the right way to play a certain role. Unlike Wolterstorff's paradigm oughts, these oughts are not based on what's normal or expected. They are based on what's good performance. Furthermore, it is plausible to say that the role of a believer is not one that we have any real choice about taking on. It is our plight to be believers. We ought to do it right. It doesn't matter that in some cases we are unable to do so.' (Feldman 2000:676)

Feldman's proposal as presented in that paper, unfortunately, is too sketchy to be properly assessed. Nevertheless, I think the underlying idea is worth considering, namely, the idea that 'role oughts' are grounded in the relevant role one plays and *all it takes* for that ground to generate 'role oughts' is for one to play that role, regardless of whether that obligation can be implemented or satisfied. Now, if doxastic oughts belong to such 'role oughts', then $\mathcal{T}_{\mathrm{O}}$ is not subject to the constraint of OIC or OICS, since the fact that S is an epistemic agent alone is sufficient to generate doxastic obligations.

---

[18] The problem is roughly this. Doxastic involuntarism is the claim that belief-formation is not under our voluntary control. Epistemic deontology is the claim that there are doxastic obligations. Given OIC, the claim that there are doxastic obligations implies that we have voluntary control over our belief-formation. Therefore, OIC, doxastic involuntarism, and epistemic deontology are jointly incompatible. For a classic discussion on doxastic involuntarism, see, Alston (1988).

However, why should we accept the claim that there are genuine 'role oughts' in the first place? For sure, we can describe what counts as a good performance appropriate to the role in question, but it does not follow that these claims issue in genuine normative requirement on the person who plays that role. Are there 'role oughts' in virtue of being a dictator, a mafiosi or a neo-Nazi? I think not. What, then, are the right criteria to determine which 'role oughts' are genuinely normative? The notion of the role of a believer seems extremely vague to me. Are there genuine 'role oughts' in virtue of being a believer? Explaining how epistemic normativity can be grounded in the role we play as epistemic agents is not a straightforward task, and it is not clear to me this is a task proponents of the invariantist accounts the truth norm would want to commit themselves to.

Now, even if we set this difficulty aside, I see no reason why 'role oughts' are not subject to the constraint of OIC or OICS. Feldman seems to conflate two senses of abilities in the passage we cited above: the basic ability to φ and the ability to φ competently. OIC is not equivalent to the claim that 'ought to φ' implies 'can φ well'. It is one thing to say that there are teachers, parents, cyclists who *cannot competently do* what they ought to do, it is another to say that there are teachers, parents, cyclists who *cannot do* what they ought to do. The former is obviously true, but the latter seems false. If one cannot implement or satisfy one's 'role oughts', is it still reasonable to count that person as playing that role? Imagine a teacher and a mother, who had an accident and lost her consciousness. It seems clear that she is not subject to the relevant 'role oughts' precisely because she lost the capacities required to carry out her role as a teacher and parent. The conditions for generating 'role oughts', whatever they are, cannot be met if one cannot implement or satisfy one's 'role oughts' since they cannot be counted as playing that role in the first place. So even if there are genuine 'role oughts', they are also constrained by OIC and OICS. Thus, I do not think appealing to Feldman's account of 'role oughts' would help proponents of the invariantist accounts of the truth norm.

Let us now turn to a second proposal. Chrisman (2008) suggests that doxastic obligations should be understood as a kind of 'rules of criticism' as opposed to 'rules of action', a distinction first drawn by Sellars (1969). *Rules of action* specify what an agent ought to do while *rules of criticism* describe how things ought to be. Only the latter but not the former are constrained by OIC and OICS. *Rules of criticism* have the following logical form:

> 'Xs ought to be in state φ, whenever such and such is the case' (Sellars 1969:508) .

We might understand *rules of criticism* as a species of evaluation describing a standard, an ideal or some desirable states of affairs, by which we can make evaluative judgment about how well Xs are

doing. As such*, rules of criticism* do not presuppose that Xs themselves can implement or satisfy the obligation to be in state φ.

Chrisman's suggestion, then, is that doxastic obligations should not be treated as *rules of action* but instead as *rules of criticism*. We are familiar with ought statements applied to inanimate objects such as 'the traffic light ought to switch from red to green every two minutes' or 'your blood pressure ought to be lower'. In such cases, it is clear that these statements are instances of *rules of criticism* and the ought-limiting principle such as OIC does not seem to apply to the subjects of those statements, after all, they are inanimate objects. When the subject of an 'ought' statement is an agent, the logical form of the statement can be ambiguous and therefore confusion may arise since we can both be in a state φ and perform an act φ. Although 'ought to believe' in its grammatical form resembles 'ought to do', that 'believe' is a verb applied to agents does not entail that doxastic oughts are *rules of action*. Instead, Chrisman argues, doxastic oughts are *rules of criticism*:

> 'The idea would be to treat doxastic oughts as a species of rules of criticism (one of perhaps several distinctively epistemic species). They are of the form X ought to be in doxastic attitude A towards proposition p under conditions C.' (Chrisman 2008:364)

If we adopt Chrisman's account of doxastic oughts as *rules of criticism*, then, instead of $\mathcal{T}_O$ and $\mathcal{F}_O$, we have the following claims:

> ($\mathcal{T}_O$*) For any S, p: S ought to be in the state of believing p when p is true.
> ($\mathcal{F}_O$*) For any S, p: S ought to be in the state of refraining from believing p when p is false.

As the examples of the traffic light and blood pressure illustrate, *rules of criticism* are not subject to OIC or OICS. So $\mathcal{T}_O$* and $\mathcal{F}_O$* are not subject to OIC or OICS.

However, I think it provides no comfort for invariantists if truth of a proposition makes a requiring claim on us in the same way as the traffic light or blood pressure is 'required' to be in a certain state. We are interested in *what claim truth can make on us as agents* with respect to our doxastic attitudes.

In a discussion that might be an anticipation of this line of objection, Chrisman emphasizes that there are essential connections between *rules of criticism* and *rules of action*. He suggests three possible accounts of the logical relationship between *rules of criticism* and *rules of action* (2008:362):

1. Conditional View: 'X ought to be φ' implies that 'if someone is responsible for X's being φ, then that person ought to do what he/she can (ceteris paribus) to bring it about that X is φ.'
2. Universal View: 'X ought to be φ' implies that 'everyone ought to do what he/she can (ceteris paribus) to bring it about that X is φ'.
3. Existential View: 'X ought to be φ' implies that 'Someone ought to do what he/she can (ceteris paribus) to bring it about that X is φ.'

Chrisman (2008:369-370) gestures towards (but does not further elaborate on) the following version of the Existential View about the logical relationship between doxastic oughts and their corresponding 'ought-to-do':

> 'S ought to be in φ' implies that 'S's epistemic community (which includes S) ought to do what they can (ceteris paribus) to bring it about that S is in φ'.

Applying this schema to $\mathcal{T}_O$* and $\mathcal{F}_O$*, we arrive at the following two claims about ought-to-do:

> ($\mathcal{T}_O$+) For every S, p: S's epistemic community (which includes S) ought to do what they can (ceteris paribus) to bring it about that S believes p when p is true.
> ($\mathcal{F}_O$+) For every S, p: S's epistemic community (which includes S) ought to do what they can (ceteris paribus) to bring it about that S refrains from believing p when p is false.

This, I think, is a quite interesting result since the corresponding ought-to-do have three distinctive features. Firstly, $\mathcal{T}_O$+ and $\mathcal{F}_O$+ are communal obligations. The truth of p not only makes a claim on an individual epistemic agent S, but also makes a claim on S's epistemic community at large. This is in contrast with $\mathcal{T}_O$ and $\mathcal{F}_O$, which apply to each individual epistemic agent. Secondly, $\mathcal{T}_O$+ and $\mathcal{F}_O$+ involve *non-doxastic* obligations such as doing what one can to provide S with evidence, arguments and institutional care (Chrisman 2008:369). And thirdly, $\mathcal{T}_O$+ and $\mathcal{F}_O$+ have ought-limiting principles built in – what one ought to do is determined by what one can do to discharge the obligation.

Thus, if doxastic oughts are *rules of criticism* along the line Chrisman suggests, the argument from Excessive Demands, Blindspot Propositions and Conflicting Norms seem to lose their grip. On the one hand, $\mathcal{T}_O$* and $\mathcal{F}_O$* are not constrained by OIC or OICS as *rules of criticism*. On the

other hand, although $\mathcal{T}_O$+ and $\mathcal{F}_O$+ as *rules of action* are subject to OIC and OICS, since they have those principles built in, they are necessarily compatible with OIC and OICS.

I think Chrisman's proposal deserves some credit. In particular, he rightly draws our attention to a relatively unexplored, but attractive idea that there are essential connections between communal, non-doxastic obligations (such as $\mathcal{T}_O$+ and $\mathcal{F}_O$+) and doxastic obligations. Having said that, however, I do not think Chrisman's proposal can be made to work for the invariantist accounts of the truth norm.

Firstly, one might challenge the relevance of a standard of evaluation 'X ought to be φ' when it is impossible to bring it about that X is in φ. There is little reason to set some impossible states of affairs as an evaluative standard by which we judge about how well X is doing. In the case of doxastic oughts, since it is impossible for S's epistemic community (which includes S) to bring it about that S believes every truth and refrains from believing every falsehood, why should we think $\mathcal{T}_O$* and $\mathcal{F}_O$*have any bearing on us?[19] Christensen (2008:5) summarizes this line of criticism nicely: 'The fact that we humans have the particular limitations we do, it is urged, is not just some trivial footnote to epistemology; it's a central aspect of our epistemic predicament. Interesting epistemology – epistemology for humans – must take account of this fact'. Let me be clear, my point is not to deny that epistemic ideal has a place in epistemology, rather, the point is to question the relevance of an ideal that is completely beyond human reach. An ideal that takes account of human limitations seems preferable as a standard of evaluation. In the absence of further arguments, there seems no reason to endorse an ideal abstracting away from human limitations.

Even if an ideal abstracting away from human limitations has a place in our epistemology, we might still question whether it is a genuine epistemic ideal, i.e., whether $\mathcal{T}_O$* and $\mathcal{F}_O$* are true. Philosophers who are sympathetic to the evaluative account of the truth norm might naturally find it compelling to accept $\mathcal{T}_O$* and $\mathcal{F}_O$* since believing all truths and avoid all falsehoods is the state of affairs with maximal epistemic value (e.g., Fassio 2011; Goldman 1999; McHugh 2012). In Chapter 2, however, I have argued against the evaluative account of the truth norm, and so, we can leave this line of response aside.

More pressingly, on Chrisman's proposal, the truth of p makes claims on us only via *rules of action*, i.e., $\mathcal{T}_O$+ or $\mathcal{F}_O$+. The truth of p is *not* a requiring or permitting reason to believe p or to refrain from believing not-p, but only a reason to do whatever we can that are instrumental to bring it about, such as collecting evidence, reasoning well, providing institutional care and so on.

---

[19] This is essentially the point I made against the evaluative account of the truth norm in Chapter 2.

This result strikes me as throwing the baby out with the bathwater, since it implies that there isn't any direct normative relation between truth and belief, a radical departure from the original view invariantists are committed to.

Therefore, appealing to Chrisman's account of doxastic oughts does not help invariantist accounts of the truth norm. Since on the one hand, understanding $\mathcal{T}_{O}$* and $\mathcal{F}_{O}$* as rules of criticism does not imply that $\mathcal{T}_{O}$* and $\mathcal{F}_{O}$* are unconstrained by OIC or OICS. On the other hand, endorsing Chrisman's account of doxastic oughts implies that truth can only make claims on us *non-doxastically*, along the line of $\mathcal{T}_{O}$+ or $\mathcal{F}_{O}$+, which amounts to a radical shift in view about the normative relation between truth and belief.

Thus, there is no easy fix of the problems facing the invariantist accounts. The invariantists cannot make those problems go away by appealing to the two prominent alternative theories of doxastic oughts we considered in this section.

6.4 Doubts about 'Ought' Implies 'Can'

Another natural move invariantists could take is to reject OIC and OICS. If epistemic obligations are not constrained by what one can believe or by whether they can be satisfied, then the problems of Excessive Demands, Blindspot Propositions and Conflicting Norms do not arise. While to the best of my knowledge OICS has not received much critical discussion, OIC has come under serious attack. So, at least with respect to Excessive Demands and Conflicting Norms, one might defend invariantist accounts of the truth norm by appealing to recent arguments against OIC. In this section, I will respond to two objections invariantists may raise against OIC: (1) objection from counterexamples; and (2) objection from empirical evidence.

Let me begin our discussion with a prominent, putative counterexample to OIC attributed to Harry Frankfurt (1969), since variants of this case has received much critical attention in recent discussion, and more importantly, it will help us to better understand the concepts of 'ought' and 'can' involved in OIC.[20] Consider the following Frankfurt-style counterexample to OIC:

> (NEUROSURGEON) Black is a neurosurgeon who can manipulate Jones' brain to ensure that Jones would act as Black wants, i.e., to kill Smith. Black prefers not to show his hands unnecessarily. He observes Jones and will only intervene to ensure

[20] See, e.g., Copp (1997); Vranas (2007); Graham (2011); Mizrahi (2012); Wedgwood (2013b, 2017).

that Jones kills Smith if Jones decides to drop his gun. Without Black's intervention Jones kills Smith.

NEUROSURGEON seems to be a counterexample to OIC. Jones could not have not killed Smith, since if Jones had decided not to kill Smith then Black would have intervened and caused Jones to kill Smith. Since Jones could not have not killed Smith, Jones cannot refrain from killing Smith. But Jones ought to refrain from killing Smith. Hence, our intuitive verdict of the case contradicts OIC: it is not the case that, necessarily, if Jones ought to refrain from killing Smith then Jones can refrain from killing Smith.

One response to NEUROSURGEON is to insist that Jones does not have a moral obligation not to kill Smith in the presence of a counterfactual manipulator such as Black (see, e.g., Vranas 2007). We should distinguish between what one morally ought to do, on the one hand, and what one is morally blameless to do on the other hand. Jones is morally blameworthy for killing Smith, but he does not in fact have an obligation to refrain from killing in the presence of a counterfactual manipulator. But why should the presence of a counterfactual manipulator make a difference when that manipulator does not show their hands as in the case of NEUROSURGEON? It seems to me wrong to say that Jones is not obliged to refrain from killing Smith in NEUROSURGEON.

More recently, Wedgwood (2013b, 2017) offered an interesting, alternative solution to the problem which respects our intuition that Jones should be held responsible in NEUROSURGEON. He suggests that although Black deprives Jones of the opportunity of not killing Smith, Black does not deprive Jones of the opportunity of not killing Smith *for reasons of his own*. There is a possibility in which Jones could have killed Smith 'differently', so to speak. The idea is that, as I understand it, the relevant ways in which one φ-s are finely individuated by one's reasons for φ-ing. Jones has the ability and opportunity to not to kill Smith *for reasons of his own,* for had Jones chosen not to kill Smith for reasons of his own, then Black would have had to show his hands forcing Jones to kill Smith. Since Jones ought not to kill Smith *for reasons of his own* and he can refrain from killing Smith *for reasons of his own*, NEUROSURGEON vindicates OIC.

Although this proposal provides an intuitively plausible explanation of NEUROSURGEON, it comes with a price: it implies a controversial view, according to which obligations are finely individuated by one's reasons for φ-ing. On a natural interpretation of this view, there is no obligation to φ *simpliciter* but only an obligation to φ *for some (right) reason R* (or an obligation *not* to φ *for some (wrong) reason R*). To properly engage with this aspect of Wedgwood's proposal would take us too far, but I will flag two potential worries for this view.

First, there is an important distinction between φ-ing in virtue of one's obligation and φ-ing merely in conformity with one's obligation since, intuitively, only in the former case one's φ-ing is not only right but also creditworthy.[21] If obligation is finely-individuated by the reason for which one φ-s as Wedgwood suggests, then it seems impossible to conform to one's obligation without doing it in a creditworthy way, since conforming to one's obligation just is to act for the right reasons. As a result, the distinction between right and creditworthy actions collapses on Wedgwood's finely individuated account of obligation.

A second worry concerns the existence of second-order reasons. A second-order reason is a reason to φ based on a first-order reason, i.e., a consideration favours φ-ing. An obligation to φ for some reason R implies that there is a reason R* to φ on the basis of R. Now, since R is a first-order reason that favours φ-ing, R* is a second-order reason according to the characterization of second-order reasons. However, it is highly controversial whether there are second-order reasons. Whiting (2017) has argued that second-order reasons cannot exist since, first, it is impossible to respond to and be guided by second-order reasons. And second, second-order reasons seem to be reasons of the wrong kind: since they are not themselves considerations that pertain to φ-ing but to the ways in which one φ-s, they are the wrong kind of considerations to be taken into account when one deliberates about whether to φ or not.[22] I do not intend to engage in this debate here; my point is simply this. *If* there are no second-order reasons, as Whiting has argued, then Wedgwood's response to NEUROSURGEON cannot be made to work since it implies the existence of second-order reasons.[23]

I think there is an alternative and simpler response to NEUROSURGEON once we get clearer about the notion of 'can' involved in OIC. Recall that according to the standard interpretation of 'can φ', S can φ just in case S has the ability to φ and the opportunity to exercise that ability.[24] So to answer the question 'can Jones refrain from killing Smith?', two considerations are relevant: (1) whether Jones has the ability to refrain from killing Smith; and (2) whether Jones has the opportunity to exercise that ability.

---

[21] See also our previous discussion on the significance of this distinction in Chapter 4, Sect. 4.3.
[22] One might question whether the distinction between first-order and second-order reasons is a deep one, in the ontological sense, given that reasons are, facts.
[23] None of the considerations are meant to be knockdown objections to Wedgwood's proposal. Indeed, I welcome further suggestions as to how Wedgwood's proposal can be made to work considering the worries we just discussed.
[24] See, e.g., Copp (2008:67 fn.2); Graham (2011:342); Mizrahi (2009, 2012); Vranus (2007:170); Wedgwood (2013b).

One immediate problem we face is how to make sense of the ability to refrain from φ-ing. Exercise of one's abilities is usually understood in the context of actions but not in the context of inactions or omissions. What does the ability to refrain from φ-ing consist of?

One natural idea is that 'refraining from killing' can be understood as involving actions on the agent's part that would avert an otherwise deadly act. Now for the ease of discussion, let us assume that in NEUROSURGEON, refraining from killing Smith involves Jones' dropping his gun. Since it is quite clear that Jones has the ability to drop his gun, the point of dispute is whether Jones has the opportunity to do so. Proponents of the counterexample believe that Jones lacks such opportunity, after all, once Black notices that Jones intends to drop his gun, he will intervene and make sure that Jones pulls the trigger and kill Smith. There is no possible world in which Jones drops his gun.

What, however, does it take for an agent to have an opportunity to φ? Wedgwood (2013b:87) offers the following analysis of opportunity in terms of objective chance:

> 'An agent x has an opportunity to *φ* at *t* at a world *w** if and only if the relevant psychological laws assign a non-zero conditional chance to the proposition that x *φ*-s at *t,* with respect to *t* and *w** – conditional on the conjunction of all propositions that are true at *w** and causally independent of how x exercises these capacities at *t*.' (2013b:87)

By psychological laws, Wedgwood means laws that are broadly based on folk-psychology and deal with our actions and attitudes. On this proposal, the relevant psychological laws assigning a non-zero chance that x will φ at time t is a necessary and sufficient condition for x to have an opportunity to φ at t.

What I find instructive in Wedgwood's analysis is that an opportunity to *φ* must be indexed to a time. Once we add time indices, it becomes clear that Jones had the opportunity to drop his gun, and therefore can refrain from killing Smith at times that are crucial to our assessment of the case.

Let t_0 be the moment when Jones made up his mind about whether to drop the gun or to pull the trigger. Let t_1 denotes the hypothetical moment when Black intervenes. From the setup of the case, we know that t_1 is later than t_0, since Black would intervene only *after* Black observes that Jones has made up his mind to drop the gun. Granted that the psychological laws assign a zero conditional chance to the proposition 'Jones drops his gun *at the time when Black intervenes*'. However, it is important to note that there is no reason to hold that the conditional chance assigned by the

relevant psychological laws to the proposition 'Jones drops his gun at t_0' is zero, since t_0 is prior to t_1 and prior to Black's intervention. Jones had the opportunity to drop his gun at t_0.

Thus, although there is no possible world in which Jones drops his gun, it does not entail that Jones did not have the opportunity to refrain from killing Smith. On our more careful analysis of 'opportunity', Jones had the opportunity to refrain from killing Smith at t_0 but if it were the case that Black intervenes, then Jones would not have the opportunity to refrain from killing Smith at t_1. This seems to me to be the right result. Jones ought to refrain from killing Smith at t_0, and Jones can refrain killing Smith at t_0. Jones is morally obliged to refrain from killing Smith at t_0, regardless whether Black would intervene later at t_1. If Black did intervene at t_1, then Jones cannot refrain from killing Smith at t_1, but Jones is not obliged to refrain from killing Smith at t_1 since the concept of obligation can no longer apply to Jones at t_1. At the time when Black intervenes Jones loses his agency and is not under any moral obligation. So, either way, OIC holds. The Frankfurt-style example is not really a counterexample to OIC.

We have seen in our discussion of the Frankfurst-style case that how we understand the concepts of 'ought' and 'can' involved in OIC is central to the debate. This leads us to another interesting, recent argument against OIC.

Proponents of OIC typically take the principle as an analytic or conceptual truth and appeal to its intuitive plausibility as ground for its justification. The claim that OIC is an analytic or conceptual truth has recently been challenged by Henne et al. (2016) on empirical grounds. The guiding thought is that if OIC is an analytic or conceptual truth, then a competent speaker should consistently judge that an agent S ought to φ when S cannot φ. We can therefore design experiments to test whether OIC is true analytically or conceptually:

> (TEST) 'If one claim analytically or conceptually entails another, then competent speakers in good epistemic positions should consistently deny the first when they know that the second is false (2016:284).'[25]

So, if a large number of competent speakers who deny that S ought to φ when they understand that S cannot φ, then we have strong evidence that OIC is not true analytically or conceptually.

Henne et al. (2016) argue that the results from recent empirical studies conducted by Chituc et al. (2016) and Buckwalter & Turri (2015) show just that – competent speakers often make

[25] Similar analysis of conceptual truth is endorsed by, for instance, Williamson (2007:74): 'A sentence s is analytic just in case, necessarily, whoever understands s assents to s.' On this analysis, the fact that a large number of competent speakers do not assent to the sentence that 'S ought to φ implies that S can φ' counts as strong evidence that OIC is not an analytic or conceptual truth.

statements contrary to OIC (given that a significant number of participants assert that S ought to φ even if they understand that S cannot φ), and therefore OIC fails to pass TEST. One might question whether the participants' judgments about OIC were sincerely made in accordance with and reflected their understanding of the concepts of 'ought' and 'can'. To this Henne et al. (2016) stress two points in reply. First, they emphasize that a significant number of participants made judgments contrary to OIC consistently in a variety of experiments conducted in Buckwalter & Turri (2015) and Chituc et al. (2016), with alternations in conditions manipulating various elements including blameworthiness, moral vocabulary, kinds of obligations, scope and durations of inability and so on. Secondly, in Chituc et al. (2016), apart from rating how much they agree or disagree a given statement, the participants also explained their answers intelligibly and demonstrated their understanding of the relevant situations. Henne et al. conclude that these empirical studies strongly indicate that OIC is not true analytically or conceptually.[26]

Granted that TEST is intuitively plausible. It follows the broadly Freagean concept of analyticity as a linguistic notion – something true in virtue of meaning.[27] However, it is far from obvious that survey-based empirical studies are the right methodology to find out whether OIC satisfies TEST. The survey-based experiments are designed to collect data about people's intuitive judgments about OIC, which tend to be fast and unreflective.

No doubt, efforts must have been made in designing those surveys to ensure that the participants do not make naïve or casual mistakes regarding the concepts of 'ought' or 'can', and to ensure that their judgments are sincerely made in accordance with their understanding of the concepts in question. Nevertheless, the data are too crude to be drawn on as evidence that OIC fails to satisfy TEST. As we have seen in our discussion of the Frankfurt-style case, the concept of 'ought' and 'can' are complex and philosophically nuanced. A philosophically untrained subject

---

[26] Chituc et al. (2016) investigate the relationship among judgments of 'ought' 'can' and 'blame'. In one of the experiments, for instance, participants were asked to read two vignettes adapted from Sinnott-Armstrong (1984) about Adams who promised to meet Brown at noon for lunch but cannot either because (low blame) his car broke down or (high blame) he left his place too late (it takes 30 minutes to drive to the restaurant). Participants were then asked whether they agree that at 11:45 a.m., Adam still ought to meet Brown. Participants were asked to respond with a number from +50 (completely agree) to -50 completely disagree and were asked to explain their answer. The results showed that a considerable number of participants judged that Adam still ought to meet Brown at 11:45 a.m. (31% in the low blame and 60% in the high blame condition).

[27] Of course, some philosophers who are sceptical about the notion of analyticity, notably Quine (1951). But this does not undermine the argument against OIC for if analyticity is untenable then OIC cannot be true analytically. One might challenge the role of linguistic and conceptual competence in determining analytic statements. Perhaps analytic truth does not imply that someone with linguistic and conceptual competence and understands the statement must assent to it. But this amounts to a shake-off the linguistic or conceptual understanding of analyticity.

is unlikely to reflect and heed the distinction between concepts involved in OIC and the other concepts in the vicinity. So it is doubtful that the participants of the experiments truly grasp the concepts of 'ought' and 'can' involved in the principle OIC and make judgments in accordance with their understanding of those concepts.[28]

To put this complaint slightly differently, it is important to notice that some conceptual entailments are much more complex than others. With respect to simple conceptual truths such as 'all bachelors are males', there is no controversy about what 'bachelors' or 'males' means, so conducting a survey would be an appropriate way to run TEST. By contrast, our pre-theoretical understanding of 'ought' and 'can' is messy. The thesis that OIC is a conceptual truth presupposes certain philosophical understanding (depending on the version of OIC) of 'ought' and 'can', which should be distinguished from our pre-theoretical, intuitive and folk-conception of 'ought' and 'can'. The empirical studies Henne et al. (2016) rely on, are simply not the relevant data for assessing the analyticity of OIC in the sense that concerns philosophers in the debate. These data are at best data about the analyticity of our pre-theoretical understanding of OIC and do not amount to an empirical refutation of OIC.

Thus, the invariantist attempt to make the problem of Excessive Demands and Conflicting Norms go away by appealing to the two recent objections to OIC fails. OIC remains a highly plausible principle.[29]

6.5 Conclusion

To sum up. In this chapter, I considered two leading responses to the weighting challenge: the obligatory and permissible account of the truth norm. I discussed three arguments against the obligatory account of the truth norm and showed that the permissible account of the truth norm faces analogous problems. I offered a diagnosis: on the reason-based normative framework, both views are invariantist, for they claim that the T-practice-based reason has a certain weight across all situations. But the T-practice-based does not always have the same requiring/permitting weight across all situations. The weight of the T-practice-based reason may vary depending on the circumstances. I then discussed two possible moves invariantists may make in response and

[28] This so-called expertise defense is defended and developed by, e.g., Williamson (2007, 2011).
[29] As we will see in Chapter 7, on the reason-based framework, at least with respect to practice-based reasons, OIC holds because if one cannot φ when R, then R as a practice-based reason for one to φ is disabled.

rejected both. In the next chapter, I shall develop and defend the variantist account of the truth norm in more detail and argue that it offers an attractive response to the weighting challenge.

7. Defending a Variantist Account of the Truth Norm

In the previous chapter, I argued that the two leading invariantist views about the truth norm face counterexamples and motivated an alternative view – variantism about the truth norm. The view, as I've stated so far, simply says that whether the T-practice-based reason is a requiring or permitting reason for belief depends on the circumstances.

In this chapter, I will develop a version of variantism by spelling out in detail just how the weight of the T-practice-based reason may vary and on what sort of features such variance may depend. I will begin by tapping into a widespread idea in contemporary metaethics which says that reasons are holistic and context-sensitive. I will suggest that a natural way to spell out variantism is on the holistic framework (Sect. 7.1). I will then address three challenges to variantism and along the way I will further develop variantism in more detail (Sect. 7.2). I will argue that variantism enables us to explain the three kinds of cases which cause trouble for the two leading invariantist views in a plausible and attractive way (Sect. 7.3). Finally, I will strengthen the case for variantism by contrasting it favourably with two alternative views. Variantism offers an attractive response to the weighting challenge (Sect. 7.4).

7.1 From Holism to Variantism

Variantism says that whether the T-practice-based reason is a requiring or permitting reason for belief can vary depending on the circumstances. In the previous chapter, we introduced the requiring/permitting distinction as a distinction about the weight of a reason, namely, whether a reason favours φ-ing is weighty enough to require or permit one to φ. The distinction is illustrated by examples and it appeals to our intuitive judgments as a guide to whether the reason for φ-ing is weighty enough to require or permit one to φ. Recall the four examples:

1. Climate emergency is an existential threat to humanity. That seems to permit me to participate in civil disobedience, but not to require it.
2. My uncle is allergic to nuts. That seems to require him to refrain from eating nuts, not merely to permit him from doing so.

3. Alice has been in contact with a Covid-19 patient and she has a fever. That seems to permit me to believe that Alice has Covid-19, but not to require it.
4. Later Alice tested positive for Covid-19 and the test was accurate. That seems to require me to believe that she has Covid-19, not merely to permit it.

Since the pioneering work by Dancy (2004), however, it is widely acknowledged that the weight of reasons is holistic and context-sensitive.[1] A consideration for φ-ing can, in one context have a certain weight, and in another context a different one.[2] For example, intuitively, the fact that climate emergency is an existential threat to humanity in some contexts can be a requiring, not merely a permitting reason for me to engage in civil disobedience. For suppose that the window of opportunity to avoid irreversible catastrophic climate consequences is closing, it may be argued that, in this context, my reason to engage in civil disobedience has a requiring weight, not merely a permitting weight.

What are the relevant factors that affect the weight of a reason in a context and is there a principled way to describe them? Following Dancy (2004), we can distinguish between two types of features that affect weights of reasons in a given context: conditions and modifiers.

Conditions can be divided into enablers and disablers. An enabler is a condition whose presence allows R to constitute a reason favouring φ-ing without itself being a reason to φ. For example, the fact that the lighting condition in my room is normal is an enabling condition for the fact that the cup appears red in front of me to constitute a reason for me to believe that the cup is red. In the presence of an enabler, R constitutes a reason for φ-ing and provides normative support for φ-ing. When the enabler goes from obtaining to not obtaining, R ceases to constitute a reason and no longer provides normative support for φ-ing. A disabler, by contrast, is a condition whose presence prevents R from constituting a reason that favours φ-ing. For example, the fact that the doctor lied to me prevents the doctor's testimony about Alice from being a reason for me to believe that Alice has Covid-

[1] Let me clear that the holistic framework I present in this chapter only draws on aspects of Dancy's work. Dancy's original holistic theory of reason is presented as a thesis that mainly concerns the *valence* of reasons.

[2] The phenomenon I try to describe here is broader than what has been widely discussed under the heading of 'contextualism' in epistemology. Epistemological contextualism is the view that the acceptability or truth-conditions of knowledge/reason ascriptions vary from context to context, where the relevant sense of 'context' is often more narrowly construed as the conversational contexts in which those knowledge/reason ascriptions are made. Here 'context' is understood in a much broader sense, referring to the circumstances where the agent finds herself in. Throughout this chapter I use 'contexts' interchangeably with 'circumstances' and 'situations'.

19. In the absence of disablers, R constitutes a reason for φ-ing and provides normative support for φ-ing. When disablers go from not obtaining to obtaining, R ceases to constitute a reason and does not provide normative support for φ-ing.

Modifiers also come in two types: intensifiers and attenuators. An intensifier is a consideration which strengthens the normative support R has for φ-ing. For example, the fact that the window of opportunity to avoid climate catastrophes is fast shrinking intensifies my reason to engage in civil disobedience. An attenuator, by contrast, is a consideration which weakens the normative support R has for φ-ing. For example, the fact that Lidl often runs out of stock attenuates the weight of my reason to go to Lidl to buy grocery.

What I said so far is familiar and widely accepted. The normative support R has for φ-ing is determined by conditions and modifiers along the lines I just described.[3] But this is not the entire story, for whether R has a requiring/permitting weight not only depends on the extent to which it supports φ-ing, but also depends on the balance of reasons in a given situation.

Let us distinguish between the degree of normative support R provides for φ-ing and the weight R has in making a requiring/permitting claim on us regarding φ-ing, I will refer to the former as the *strength* of R and the latter the *weight* of R. The *strength* of R is subject to conditions and modifiers, while the strength of R and its balance against the strength of its opposing reasons determine the *weight* of R in making a requiring/permitting claim on us regarding φ-ing.

Other things being equal, when reasons against φ-ing have more strength, R is less weighty in making a claim on the agent regarding φ-ing, and if they have less strength, R is more weighty in making a claim on the agent regarding φ-ing. Let me elaborate.

Reasons against φ-ing, just like reasons for φ-ing, are considerations that bear upon φ-ing.[4] But unlike reasons for φ-ing, a reason against φ-ing is a fact that counts against φ-ing. For example, the fact that engaging in civil disobedience is a criminal offence is a reason against the act. And the fact that it distresses my family is another reason against engaging in

---

[3] Notice that a minority of philosophers have challenged the distinction between reasons, conditions and modifiers. See, e.g., Crisp (2000), Raz (2000) and Fogal (2016). For a recent defence of the reasons/conditions/modifiers distinction, see Bader (2016). In this chapter, I am taking it for granted that the distinction is meaningful and robust.

[4] Alternative views about *reasons against* holds that *reasons against* are reduced to *reasons for*, for example, φ-ing are reasons for not φ-ing (e.g., Nagel 1970:47; Schroeder 2007: Chp.7). I will set this debate aside. For further discussion on this issue, see Snedegar (2018).

civil disobedience. It is easy to see that *reasons against* are also subject to conditions and modifications in the same way as *reasons for* are subject to conditions and modifications. An enabler of a reason R against φ-ing enables R to constitute a reason against φ-ing while in the presence of a disabler, R ceases to constitute a reason against φ-ing. For example, the fact that my action distresses my family ceases to be a reason against engaging in civil disobedience if I know that my family are trying to manipulate me emotionally. The strength of a reason against φ-ing can also be attenuated or intensified. The reason that civil disobedience is a criminal act has less strength in the presence of attenuators such as the fact that the penalty is lenient, and it has more strength in the presence of intensifiers such as the fact that an authoritarian regime has come into power.

We cannot fully explain our intuitive judgment about the varying weight of the same reason in different contexts unless we also include the strength of opposing reasons in our analysis. To illustrate, consider and compare the following cases:

(CASE 1) There is a climate emergency. Engaging in civil disobedience is a criminal offence and it distresses my family.

(CASE 2) There is a climate emergency. Engaging in civil disobedience is a criminal offence and it distresses my family. An authoritarian regime has come into power and imposed a much harsher penalty on civil disobedience.

(CASE 3) There is a climate emergency. Engaging in civil disobedience is a criminal offence and it distresses my family. But the punishment is lenient, and I know that my family are trying to manipulate me emotionally.

Although the normative support climate emergency provides for engaging in civil disobedience (i.e., the *strength* of the reason) remains the same across three cases, it seems clear that climate emergency carries different *weights* in making a claim on me regarding whether to engage in civil disobedience. Climate emergency in CASE 3 seems to be a more demanding reason for me to engage in civil disobedience than that in CASE 1, which is still more demanding than that in CASE 2.

On the account I just sketched, we can explain the varying weights of climate emergency in the three cases, despite having the same strength. In both CASE 1 and CASE 2 the weight of my reason to engage in civil disobedience is weakened by the presence of the two opposing

reasons, namely, the fact that engaging in civil disobedience is a criminal offence and it distresses my family. But in CASE 2 the strength of one of the opposing reasons is further intensified by the prospect of harsh punishment and thus it weakens the weight of my reason to engage in civil disobedience to a greater extent compared to that in CASE 1. In CASE 3, one of the opposing reasons present in the other two cases – the concern about my family's distress is disabled by the fact that my family are trying to manipulate me emotionally. The strength of the remaining opposing reason is further attenuated by the fact that the punishment is lenient. Thus, the weight of my reason for engaging in civil disobedience is the most demanding in CASE 3 where the opposing reasons have the least strength.

With the holistic account of reason in place, we can now characterize the distinction between requiring and permitting reason as a distinction about the weight of a reason, which is determined by both the normative strength of that reason and the normative strength of its opposing reasons in a given situation. A consideration R is a *requiring* reason for S to φ in a given situation just when R is a reason for S to φ and its normative strength *is greater than* the normative strength of its opposing reasons combined in that situation. A consideration R is a *permitting* reason for S to φ in a given situation just when R is a reason for S to φ and its normative strength is *at least as great as* the normative strength of its opposing reasons combined in that situation.[5] Thus, if the normative strength of climate emergency is greater than that of all the reasons I have against engaging in civil disobedience, then the concern about climate emergency requires me to take the action. If the normative strength of all the reasons I have against engaging in civil disobedience is as strong as the normative strength of climate emergency, then the concern about climate emergency permits me to engage in civil disobedience though it does require me to do so.

The holistic framework we just sketched provides a natural way to spell out variantism. The core claim is as follows: the T-practice-based reason to believe p and to refrain from believing not-p when p is true is holistic and context-sensitive *just like any other reasons*. Whether the T-practice-based reason has a requiring or permitting weight can vary depending on the presence/absence of conditions, modifiers and opposing reasons in a given

[5] There is a complication concerning how the normative strength of different opposing reasons should be combined – they do not always add up in a strictly additive way for some of them might overlap. This raises question about what it is for reasons to overlap. One recent suggestion is that we can distinguish between load-bearing and derivative reasons for and against an option and only the latter contribute non-overlapping normative strength to an option (e.g., Maguire & Snedegar 2021). Pursuing this issue will take us too far. For our present purposes, I will assume that our intuitive understanding of these notions l suffices for us to make reasonable judgments about the combined normative strength of opposing reasons in a situation.

situation. The truth of a proposition p is a *requiring* reason for S to believe p just in case that the truth of p is a reason for S to believe p and its normative strength is greater than the normative strength of opposing reasons against believing p in that situation. And the truth of p is a *permitting* reason for S to believe p just in case that the truth of p is a reason for S to believe p and its normative strength is *as least as great as* the normative strength of opposing reasons against believing p in that situation.

To sum up, in this section I suggested a natural way to spell out variantism by tapping into the widespread idea in contemporary metaethics which says that reasons are holistic and context-sensitive. Variantism is simply the result of applying the holistic account of reasons to T-practice-based reasons for belief. Truth as a reason for belief functions holistically just like any other reasons: its weight in making a claim on us regarding what to believe can vary depending on the presence/absence of conditions, modifiers and opposing reasons in a given situation. This view, I submit, offers a novel and attractive response to the weighting challenge, or so I shall argue in the remainder of this chapter.

7.2 Variantism: Conditions, Modifiers, and Opposing Reasons

Three question arise concerning the move from holism to variantism. First, can the T-practice-based reason be conditioned? If so, what sort of facts constitute conditions that enable/disable T-practice-based reason to believe p and refrain from believing not-p when p is true? Second, can the T-practice-based reason be modified? If so, what sort of facts constitute modifiers that intensify/attenuate T-practice-based reason to believe p and refrain from believing not-p when p is true? Finally, can truth be weighted by its opposing reasons? If so, what sort of facts constitute its opposing reasons? My goal in this section is to answer these questions, and in doing so, to further develop the variantist account of the truth norm.[6]

[6] Despite the widespread use of these notions in the literature of reason, most philosophers rely on intuitive examples to illustrate the distinctions between reasons, conditions and modifiers. Recently, Maguire (2016) offered a detailed account of the distinction based on his value-based theory of reason. Bader (2016) offered a more general metaphysical framework which allows one to establish the distinction without committing to a particular theory of reason. The account I offer below is influenced by but remains substantially different from the two accounts.

7.2.1 Can the T-practice-based Reason be Enabled/Disabled?

Variantism claims that whether the T-practice-based reason has a requiring/permitting weight partly depends on whether that reason is enabled/disabled in a given situation. Disabling conditions, as we have said, are what prevent a fact R from counting in favour of φ-ing. Enabling conditions, by contrast, are what allow R to count in favour of φ-ing. But can the T-practice-based reason be enabled/disabled? And if so, what sort of facts constitute enabling/disabling conditions of truth?

I have argued in Part I of this thesis that the authority of the truth norm is grounded in what I call the T-practice: a justified social practice of belief management governed by the truth norm of belief. I suggested that a practice-based reason R in favour of φ-ing is grounded in a canonical fact of the following form:

(Practice-based) If R were to obtain, φ-ing would constitute a justified social practice.

Practice-based reasons in general are subject to conditions that enable/disable the relevant practice-based grounding facts to obtain. There are two possible ways in which a practice-based reason can be disabled. First, a practice-based reason to φ can be disabled for S if S cannot φ, for in that case, it is not possible for S to participate in the practice which grounds the reason to φ. For example, in normal circumstances, dunking is part of basketball playing and there can be practice-based reasons to dunk in a game of basketball. But dunking on a standard-sized basketball hoop cannot be part of basketball playing for young children since it is not something young children can do. So, for young children, the practice-based reasons to dunk are disabled – they do not have any practice-based reasons to dunk when they play basketball.

Second, a practice-based reason to φ can be disabled for S if the practice which grounds the reason to φ is unjustified in a given situation. For example, in normal circumstances, playing chess is a justified social practice. And you have a practice-based reason not to move your knight horizontally. But suppose that you know your opponent will be killed if you do not move your knight horizontally. In this case, the game of chess you are engaging in is no

longer a justified social practice (for it endangers your opponent). The practice-based reason you have for not moving your knight horizontally is therefore disabled.[7]

Thus, practice-based reasons in general are not unconditional. They can be enabled/disabled by facts about what participants of that practice can or cannot do and facts about whether the practice is justified/unjustified in the situation under consideration. Truth, as a practice-based reason, is not unconditional either. The T-practice-based reason can be enabled/disabled by facts about what an ordinary epistemic agent can or cannot do with their beliefs and by facts about the justificatory status of the T-practice which grounds the T-practice-based reason in the situation under consideration.

In the previous chapter, we discussed several examples where the T-practice-based reason fails to have a requiring weight when an ordinary epistemic agent cannot believe p or refrain from believing not-p when p is true. We are now able to explain why this is so without the need to invoke the principle OIC. This is because the T-practice-based reason is *disabled* in those cases when the epistemic agent cannot believe p or refrain from believing not-p.

One example we discussed concerns complex propositions that are impossible for any ordinary epistemic agent with limited cognitive capacity to form doxastic attitudes about. The fact that we cannot form beliefs about those complex propositions disables the T-practice-based reason for us to believe those complex propositions when they are true. For in that case, it is not possible for us to participate in the T-practice with respect to those propositions which grounds the reason for us to believe them. In another example, a patient with Capgras delusion cannot refrain from believing that a close relative has been replaced by an imposter due to cognitive failure. Refraining from believing that the relative is an imposter is not something one who suffers from that medical condition can do. It is not possible for them to participate in the T-practice with respect to those propositions which grounds the relevant reasons. Therefore, the fact that one suffers from Capgras delusion and cannot refrain from believing that her relative is an imposter disables the T-practice-based reason for her to refrain from believing that proposition. In a range of more fanciful examples, we consider propositions such as that time does not flow, that $2+2\neq4$, and that I

[7] One might think that I fail to draw a distinction between reasons for/against participating in a practice and practice-based reasons for/against φ-ing. According to this line of objection, the fact that a practice is unjustified, is a reason against participating in that practice, but the practice-based reasons for/against φ-ing remain unaffected. So, the fact that the game of chess is unjustified is a reason for you not to play that game but the practice-based reason to not to move your knight horizontally remains unaffected. But this cannot be right, for reasons are authoritative and normatively-binding. When a player's safety is at risk, there is no authoritative, normatively-binding reason to refrain from moving your knight horizontally.

am a brain in a vat. Suppose that they are true. Again, in these cases, they do not constitute reasons for me to refrain from believing that time flows, that 2+2=4 and that I am not a brain in a vat, since arguably it is not something an ordinary epistemic agent who understands basic arithmetic and experiences time can do. It is not possible for us to participate in the T-practice with respect to those propositions which grounds the reason for us to refrain from believing them when they are false. Therefore, the T-practice-based reasons for me to refrain from believing that time flows, that 2+2=4, and that I am not a brain in a vat are disabled by the fact that I cannot refrain from believing those propositions.[8]

Practice-based reasons for belief can also be disabled if the T-practice which grounds practice-based reasons for belief is unjustified in a given situation. For example, believing certain things about a marginalized community, even when they are true, can often amplify an unjustified narrative about the community and put members of that community in harm's way. If believing such propositions constitute doxastic wrongdoing, as some philosophers have argued (e.g., Basu 2018, 2019), then believing those true propositions which constitute doxastic wrongdoing cannot be a justified social practice of belief management. So, arguably, the fact that believing p constitutes doxastic wrongdoing in a given situation can disable the T-practice-based reason to believe p when p is true in that situation.

In short, the T-practice-based reason for one to believe p/refrain from believing not-p can be enabled/disabled by facts about what an ordinary epistemic agent can or cannot do and by facts about the justificatory status of the T-practice in the situation under consideration.

---

[8] The set of examples of disablers concerns what an ordinary epistemic agent cannot do with respect to her beliefs in a given situation. But let me be clear, such disablers are not cheap to come by. One cannot φ in a given situation if one does not have the ability or opportunity to φ in that situation, broadly based on folk-psychology laws governing our attitudes and actions. There is a line to be drawn between things that are *extremely difficult to believe under certain conditions* and *things that are impossible to believe under those conditions*. A Trump-supporter who has been fed with lies and conspiracy theories about the presidential election *does not* have his *T-practice*-based reasons disabled. Even if it is difficult for him to come to believe the truth about the election, after all, his epistemic environment is impoverished, and he is emotionally manipulated. Nonetheless, he has fully functioning cognitive capacities to think and make correct judgments. He has the opportunity to reflect upon what is going on around him, to review the evidence and to reason. There is no reason to think that it is *impossible* for him to come to believe the truth. Of course, there might be borderline cases where the line between what one can and cannot believe is harder to draw. But those cases need to be carefully assessed on a case-by-case basis with the relevant detail filled in.

7.2.2 Can the T-practice-based Reason be Modified?

Variantism also claims that whether the T-practice-based reason has a requiring/permitting weight partly depends on its normative strength modified by intensifiers/attenuators in a given situation. Intensifiers, as we have said, are what strengthen the normative support a reason R has for φ-ing, whereas attenuators are what weakens the normative support R has for φ-ing. But can the T-practice-based reason be modified by intensifiers/attenuators? And if so, what sort of facts constitute its modifiers?

Notice first that desire-based and value-based reasons can be intensified/attenuated by facts concerning the extent to which the states of affairs constituted by φ-ing are desired or valuable, since both desired and valuable states of affairs are gradable, i.e., they can be more or less desired or valuable. For example, the fact that Davidson writes elegant prose intensifies my reason to read his work since doing so in the presence of that fact constitutes a more desired states of affairs than it would otherwise be in its absence. While the fact that it is difficult to follow Davidson's reasoning attenuates my reason to read his work since doing so in the presence of this fact constitutes a less desired states of affairs than it otherwise would be in its absence. The same can be said for value-based reasons. For example, the fact that the pedestrian is an elderly lady intensifies my reason to help her since doing so in the presence of this fact constitutes a more valuable states of affairs than it otherwise would be in its absence.

A practice is not gradable in the same way as desires or values. Either φ-ing for R would constitute a justified social practice or it would not. But plausibly, we can distinguish between the centre and the periphery of a social practice relative to its function. The distance between the centre and the periphery is gradable and φ-ing for R in a situation can be closer or less close to the centre of the practice. Practice-based reasons can be intensified/attenuated by facts concerning how central φ-ing for R is within the practice relative to its function in a given situation.

We can apply this idea to the T-practice with the help a simple diagram of concentric circles. Consider the following simplified diagram (Figure 3):

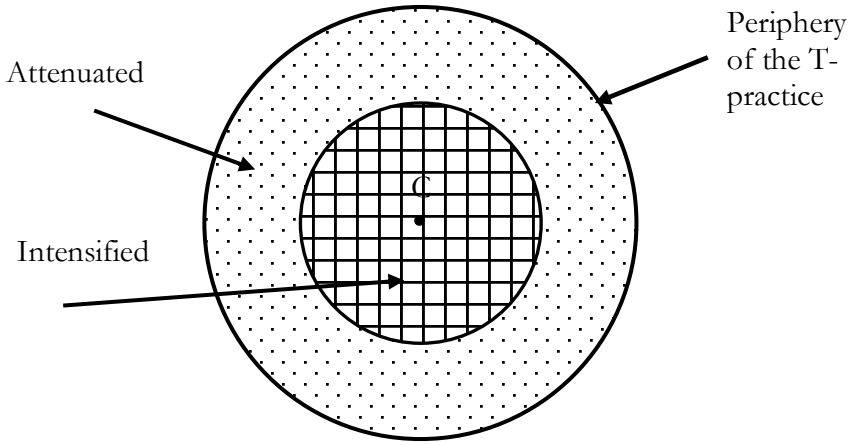


Figure 3: Practice-based Modifiers

Let C represents the centre of the T-practice. Let any point falling within the periphery of the T-practice represents a situation in which one believes p/refrains from believing not-p when p is true. We can model this type of modifiers of T-practice-based reasons using the diagram. The fact which says that believing p/refraining from believing not-p when p is true in a given situation falls within the inner circle, i.e., in the gridded area, constitutes an intensifier. The shorter the distance to C, the greater extent to which it intensifies the T-practice-based reason to believe p/refrain from believing not-p when p is true in the situation under consideration. The fact which says that believing p/refraining from believing not-p when p is true in a given situation falls outside the inner circle but within the periphery of the T-practice, i.e., in the dotted area, constitutes an attenuator. The longer the distance to C, the greater extent to which it attenuates the T-practice-based reason to believe p when p is true in the situation under consideration.[9]

This, of course, raises the question about what constitutes the centre of the T-practice. Here we can draw on the work done in Chapter 3, where I argued that the function of the T-practice is to facilitate social cooperation, knowledge production and maintenance. At a first approximation then, we can say that if believing p/refraining from believing not-p when p is true in the situation under consideration plays a greater role in facilitating social cooperation, knowledge production and maintenance, then the closer it is to the centre of

[9] Any fact which says that believing p for the truth of p in a given situation falls *on* the inner circle can be understood as a modifier which maintains the unmodified weight of the *T-practice*-based reason in that situation.

the T-practice. Of course, a good deal more can be said to make this characterization more precise.[10] But I take the general idea to be intuitive enough.

Let's apply this idea first to the example of a blindspot proposition such as 'it is rainy but nobody believes that it is rainy'. Since blindspot propositions are necessarily false if one believes them, believing a blindspot proposition p when p is true, one might think, is at the very periphery of the T-practice, since forming necessarily false beliefs does little to facilitate social cooperation, knowledge production or maintenance. The fact that p is a blindspot proposition attenuates the T-practice-based reason to believe p when p is true. By contrast, when the survival and flourishing of our community hinges on us grasping the truth about p, even if p may be painful to accept or hard to uncover, believing p when p is true is at the centre of the T-practice. For example, when p concerns issues such as the Covid-19 pandemic, climate change, social justice, etc., the T-practice-based reason is intensified. In such cases, believing p and refraining from believing not-p when p is true is crucial to facilitate social cooperation, knowledge production and maintenance.

These examples are relatively simple and straightforward. But you might think there are much more complex cases where, for instance, believing p/refraining from believing not-p might be essential to promote social cooperation but threatens knowledge production.[11] And in some cases it is difficult to make a judgment about the importance of believing p/refraining from believing not-p when p is true in a given situation relative to the function of the T-practice. Such cases would require a much more worked-out theory, a task I have to leave to future work. Again, at this stage it is important to emphasize a close case-by-case analysis. The quick sketch suffices to illustrate how the framework I just offered can model one way in which the T-practice-based reason can be modified, which is our focus here.

I will now turn to another important type of modifiers: risk attenuators. The thought is this: a reason R to φ can give an agent S normative support for φ-ing in a given situation only if R in fact constitutes a reason for S to φ in that situation. It is quite plausible to think that in our world of uncertainty, a reason R to φ in a given situation can involve certain risks. For example, the fact that a project will lift one hundred families out of poverty is a reason to invest in that project. But in our world of uncertainty, there is a risk that the project will

---

[10] For instance, what constitutes the centre/periphery of the *T-practice* may change over time and differ across different cultures and communities. Also, there may be borderline cases and there may not be sharp cut-off points between intensifier and attenuators.

[11] There is the standard challenge of how to get a proper ordering when there are multiple criteria: it is more a vector-based approach of relevance than a straightforward ordering.

not lift one hundred families from poverty. The risk that the project will not lift one hundred families out of poverty attenuates my reason to invest in that project. According to an intuitive probabilistic account of risk, the risk of R being false is determined by the probability of R conditional upon the background evidence available to the agent in a given situation. The higher the evidential probability of R, the less risky is R.[12] The risk that my reason to invest in the project does not obtain, i.e., the high probability that the project will not lift one hundred families from poverty given my evidence, attenuates my reason to invest in that project. We might call this type of risk, *empirical risk*. It is the risk of R being false in a given situation. If R were false, then R would not constitute a reason for φ-ing in that situation. The higher the risk, the weaker the normative support R has for φ-ing.

Empirical risk can be contrasted with what we might call *normative risk*. The latter is the risk that R fails to count in favour of φ-ing in a given situation, i.e., the risk of the grounding fact in virtue of which R counts in favour of φ-ing failing to obtain.[13] If R did not count for φ-ing in a given situation, then R would not constitute a reason for φ-ing in that situation. For example, the fact that a project benefits future generations is a reason for investing in that project, grounded in the value-based fact that the welfare of future generations constitutes valuable states of affairs. But there is a risk that we are wrong about how much we should care about future generations. Perhaps, one may argue, that only the welfare of the present and near-future generations matters. The normative risk that the future-oriented project fails to constitute valuable states of affairs attenuates my reason to invest in that project. According to the probabilistic account of risk, the normative risk of R failing to count for/against φ-ing is determined by the probability that the grounding fact in virtue of which R counts for φ-ing obtains, conditional upon the background evidence available to the agent in a given situation. The higher the evidential probability for the grounding fact, the less (normatively) risky is R. The higher the normative risk, the weaker the normative support R has for φ-ing.

T-practice-based reasons can be attenuated by risk attenuators. Consider first the empirical risk that p is in fact false. If p were false, p would not constitute a T-practice-based reason to believe p/refrain from believing not-p. For example, suppose that Fin was hiking

---

[12] Recent work suggests there might other plausible accounts of risk, such as the modal and normic accounts of risk. See, e.g., Pritchard (2016, 2020), Ebert et al. (2020). For the ease of presentation, I will work with the probabilistic account of risk.

[13] The idea of normative risk is gaining traction in recent work, albeit discussed under different labels such as 'moral risk' (e.g., Weatherson 2014), 'moral uncertainty' (e.g., Bykvist 2017) and 'normative uncertainty' (e.g., Pittard & Worsnip 2017).

in the Cairngorms National Park and saw a large white object. It was a snowy day and the ground was covered in white. Fin saw that the children of local farmers were out playing in that area. Fin has never seen a mountain hare in this part of the park. But luckily for Fin, the large object he just spotted was a mountain hare. In this case, the T-practice-based reason for Fin to believe that he saw a mountain hare is of a high empirical risk. For given the evidence available to Fin, the evidential probability that Fin saw a mountain hare is low. The risk attenuates Fin's T-practice-based reason to believe that he saw a mountain hare.

Consider next the normative risk that the practice-based grounding fact in virtue of which the truth of p constitutes a reason does not actually obtain in a given situation. To illustrate, consider an example adapted from Basu (2019:915-916). Suppose that the conference at Aanya's university has ended, and the participants are having dinner at a local restaurant. After a few drinks, Aanya gets up to use the restroom. As she returns to her table, one of the diners, Jim, asks her for another drink – he believes that Aanya works in the restaurant. With respect to their melanin levels, Aanya's is more similar to those of the waiting staff than to her fellow philosophers. Aanya feels hurt and is upset. But as a matter of fact, Aanya does have a part-time job in the restaurant. Jim's belief that Aanya works in the restaurant is true. In this case, although Jim has a T-practice-based reason to believe that Aanya works in the restaurant since it is true, intuitively, the T-practice-based reason is attenuated by a normative risk, or so I will suggest.

I noted earlier that the T-practice-based reason to believe p when p is true in a given situation can be disabled, if believing p constitutes epistemic wrongdoing and renders the T-practice unjustified in that situation. One example I provided, is a case where believing certain truths about a marginalized community amplifies an unjustified narrative about the community and puts members of that community in harm's way. I suggested that in such cases T-practice-based reasons to believe those propositions are disabled. Now, in the present case concerning Aanya, no such disabler is present, for we lack conclusive evidence as to whether the moral implications of Jim's belief render the T-practice unjustified. Nevertheless, given the pain Jim's belief has inflicted on Aanya, there is a normative risk that the T-practice-based reason does not in fact count in favour of believing that Aanya works in the restaurant.[14]

---

[14] Interestingly, in this case, it seems that the *empirical* risk that Aanya is *not* a member of the waiting staff is significant, despite the high evidential probability that Aanya is a member of the waiting staff rather than a fellow philosopher, based on the statistics. This sort of case could be used to motivate a non-probabilistic account of risk along the lines of Ebert et al (2020).

On the account I offer, T-practice-based reasons to believe the truth can be modified. The normative strength of a T-practice-based reason to believe p depends on how important is believing p/refraining from believing not-p when p is true relative to the function of the T-practice in facilitating knowledge production, maintenance and social cooperation. And it also depends on the empirical and normative risk of that T-practice-based reason in the situation under consideration.

7.2.3 Can the T-practice-based Reason be Weighed Against?

Variantism claims that whether truth has a requiring/permitting weight also depends on the normative strength of opposing reasons against believing p. But can truth as a reason for belief be weighted by reasons against believing p? If so, what sort of facts constitute opposing reasons that can play a role in weighting truth?

Here we face a complication. As we have already noted, there is an ongoing debate between *evidentialists* and *pragmatists* about reasons for/against belief.[15] Roughly, evidentialists hold that *only* considerations that bear upon the falsity of p, i.e., evidence indicating not-p, can constitute reasons against believing p.[16] According to pragmatists, non-evidential considerations can also constitute reasons against believing p.[17] For example, one simple putative non-evidential reason against believing p is the fact that believing p will take up cognitive resources such as storage space and processing power (e.g., Harman 1986). The fact that I have no interest in p seems to be another non-evidential reason against believing p (e.g., Papineau 2013). Interpersonal relationships and expectations may also give rise to non-evidential reasons against having certain beliefs about one another. For example, the fact that we are friends is presumably a non-evidential reason against believing less well of you (e.g., Keller 2004; Stroud 2006). The fact that believing p is morally wrong seems to be a non-evidential reason against believing p (e.g., Way 2012). The list goes on and evidentialists would need to explain these cases away.[18] Pragmatists, on the other hand, face

---

[15] One way to characterize the debate between pragmatists and evidentialists is in terms of the grounding facts of reasons for/against belief, namely, it is a dispute about whether there can be desire-based or value-based reasons for/against believing p that do not bear on the truth/falsity of that belief. Evidentialists think not, while pragmatists think there can be such reasons.
[16] See, e.g., Kelly (2002); Shah (2006); and Way (2016).
[17] See, e.g., Papineau (2013); Rinard (2015, 2017, 2019); Reisner (2018); Steglich-Petersen & Skipper (2019); and Woods & Maguire (2020).
[18] For a recent overview of such cases, see, Woods & Maguire (2020:212-213).

the charge that the putative non-evidential reasons are 'the wrong kind of reasons' and have no bearing on what one ought to or may believe.[19]

Notice that what we have said so far in this chapter in fact provides a way to explain some of those putative non-evidential reasons away. On my account, many of the putative non-evidential reasons against believing p are in fact disablers or attenuators of the *T-practice*-based reason to believe p when p is true. For example, the fact that p concerns some trivial issue or the fact that believing p is morally wrong, may disable or attenuate the T-practice-based reason to believe p when p is true in that situation. Such non-evidential considerations are normatively significant and play an important role in weighting the *T-practice*-based reason, though they are not themselves reasons against believing p. This, of course, is far from a conclusive argument against the existence of non-evidential reasons, but it does lend support for taking evidentialism as our default assumption in spelling out variantism. In what follows, therefore, I will just focus on the evidential reason E against believing p, where E is the body of evidence indicating not-p available to an epistemic agent in a given situation.[20] I will set aside putative desire-based and value-based non-evidential reasons against believing p.[21]

The *T-practice*-based reason one has for believing p when p is true has less weight in making a claim on the agent regarding whether to believe p when E has more normative strength, other things being equal. With these clarifications in place, the next question then, is how the normative strength of E is determined in a given situation.

Starting with conditions, since E can be derived from the T-practice-based reason to refrain from believing p when p is false, E is grounded in (1) the fact that E indicates that p is false and (2) the fact that p is false is a T-practice-based reason to refrain from believing p.[22] E can be disabled if either (1) or (2) fails to obtain in a given situation. For example, the

---

[19] See, for instance, Hieronymi (2005) and Reisner (2009).

[20] By E, I mean the entire body of evidence against believing p available to an agent. By focusing on the modified weight of the entire body of evidence against believing p available to an agent, as opposed to individual pieces of evidence against believing p, we can sideline the issue regarding how to combine the modified normative strength of different pieces of evidence against believing p.

[21] It doesn't mean that I'm not open to the idea that there can be non-evidential reasons for/against belief. It should be noted that variantism itself is compatible with both evidentialism and pragmatism and it offers a framework for weighting truth against evidential as well as non-evidential reasons. If there are indeed non-evidential reasons against believing p, we can add them to the plate of opposing reasons in weighting the truth.

[22] There is a further debate about what grounds evidential reasons for belief. While some pragmatists hold that evidential considerations can constitute reasons for/against belief only if they are desire-based or value-based (e.g., Papineau 2013; Rinard 2015, 2017, 2019), other pragmatists think evidential reasons for/against belief are constitutive (e.g., Woods & Maguire 2020). Since evidential

fact that the cup appears red constitutes an evidential reason against believing that the cup is green. That reason is enabled by the fact that the lighting condition is normal. It can be disabled by the fact that one suffers from green-red colour blindness, for in that case the fact that the cup appears red no longer indicates the falsity of the claim that the cup is green. E can also be disabled by facts that disable the non-derivative *T-practice*-based reason from which it derives, such as facts about what an ordinary epistemic agent cannot believe and by facts about the justificatory status of the *T-practice* in the situation under consideration as I argued in Sect. 7.2.2.

As a derivative reason, the degree of normative support E has against believing p is determined by how strongly E indicates that p is false in a given situation, which is subject to two risk attenuators. First, it is subject to the empirical risk of E being false in a given situation. If E were false, then E would not constitute a reason against believing p in that situation. On the probabilistic account of risk, the higher the evidential probability of E, the less empirically risky is E and the stronger the normative support E has against believing p. For example, although the fact that Smith has an alibi constitutes an evidential reason against believing that Smith committed the murder, there is a risk that Smith does not in fact have an alibi. Other things being equal, the higher the empirical risk that Smith does not have an alibi, the weaker is the reason against believing Smith committed the murder. Second, there is also a normative risk that E fails to count against believing p in a given situation, i.e., there is a risk that E does not in fact indicate the falsity of p in that situation. On the probabilistic account of risk, the higher the evidential probability that E constitutes evidence against believing p, the less normatively risky is E and the stronger the normative support E has against believing p. For example, the fact that Black testifies that Smith committed the murder constitutes an evidential reason against believing Smith is innocent. But there is a normative risk that Black's testimony does not constitute evidence, since Black's testimony may not be reliable. Other things being equal, the higher the normative risk that Black's testimony does not constitute evidence, the weaker is that evidential reason against believing that Smith is innocent.

Now, putting these pieces together, we can fully flesh out the variantist account of the truth norm. Whether a T-practice-based reason has a requiring/permitting weight depends on the normative strength of that T-practice-based reason and its balance against its

---

considerations are derivative from the *T-practice*-based reason to believe p and to refrain from believing not-p when p is true, evidential reasons on my view are practice-based reasons.

opposing evidential reasons in a given situation. The truth of p is a *requiring* reason for S to believe p/refrain from believing not-p in a situation just in case that the truth of p is a reason for S to believe p/refrain from believing not-p and its normative strength is greater than the normative strength of opposing evidential reasons in that situation. And the truth of p is a *permitting* reason for S to believe p/refrain from believing not-p just in case that the truth of p is a reason for S to believe p/refrain from believing not-p and its normative strength is *as least as great as* the normative strength of opposing evidential reasons in that situation.

7.3 The Appeal of Variantism

We have now answered all three questions concerning the move from holism to variantism. I have spelled out in detail how truth can be conditioned, modified and weighted by reasons against believing p. The T-practice-based reason functions holistically just like any other reasons: the weight it has in making a claim on us regarding what to believe can vary depending on the circumstances along the line I sketched. In this section, I will put variantism to work and argue that variantism enables us to explain the three problems facing invariantist accounts discussed in the previous chapter in an attractive and economical way.

In the previous chapter, we discussed three arguments against two leading invariantist accounts of the truth norm: the obligatory and the permissible account of the truth norm. We have seen three kinds of counterexamples to the two invariantist accounts:

(a) Cases involving propositions which we cannot believe/refrain from believing.
(b) Cases involving blindspot propositions.
(c) Cases involving propositions unsupported by evidence.

Variantism enables us to explain these cases in an attractive and economical way. It is attractive and economical because with the holistic toolkit, we can deliver correct verdicts and a nuanced analysis of the normative import of truth in a given case without relying on ought-limiting principles and evidential norms employed in the arguments we discussed in the previous chapter – they are simply the results of the fact that the T-practice-based reason, just like any other reasons, functions holistically and has varying weight in different situations.

It will be helpful to go through some examples. Let us begin with an (a)-type cases. Consider the following example:

> (PASSING TIME) It appears to me that time passes. Days end, months go by and years flow. It is impossible for me to refrain from believing that time passes. However, as a matter of fact, the passage of time is illusory.

In this case, that the passage of time is illusory is not a requiring reason for me to refrain from believing that time passes, as we have argued in Sect. 6.2.1, since it is incompatible with the ought-limiting principle OIC. Given variantism, however, we can explain why this is the case without the need to invoke OIC. In PASSING TIME, the presence of a disabler, namely, the fact that it is impossible for me to refrain from believing that time passes prevents the truth to constitute a practice-based reason for me to refrain from believing that time passes and therefore the T-practice-based reason in PASSING TIME does not have any weight in making a claim on me. This explanation generalizes to other type (a) cases where the T-practice-based reason to believe p/refrain from believing not-p when p is true is disabled. The fact that the proposition involved is something one simply cannot believe or refrain from believing disables the truth of that proposition from constituting a practice-based reason and hence it makes no claim on us regarding what to believe/refrain from believing.

Type (b) cases involve blindspot propositions. Consider the following example:

> (RAIN) Ginger is at a conference with a group of philosophers. The talk is engaging, and the discussion is animated. Suppose that, as a matter of fact, it is rainy outside and no one in the room believes that it is rainy outside.

In this case, the T-practice-based reason to believe that blindspot proposition – it is rainy outside and no one in the room believes that it is rainy outside – cannot be a requiring reason, as we have argued in Sect. 6.1, since it is incompatible with OICS. Nor is it a permitting reason, as we have argued in Sect. 6.2, since a permission to believe a true blindspot proposition is normatively 'fishy'. Again, variantism enables us to explain why in this case, the T-practice-based reason is neither a requiring nor a permitting reason without the need to invoke OICS or the 'fishiness' of the permission to believe a blindspot proposition.

According to variantism, the T-practice-based reason has neither a requiring weight nor a permitting weight because the normative strength of the T-practice-based reason to believe the true blindspot proposition in RAIN is attenuated and outweighed by opposing evidential reason against believing that blindspot proposition. As I have argued, the fact that believing the blindspot proposition – it is rainy outside and no one in the room believes that it is rainy outside – is at the very periphery of the T-practice attenuates the normative strength of the

T-practice-based reason to believe that blindspot proposition. Moreover, there is also an empirical risk that the T-practice-based reason in RAIN does not obtain. After all, even if the talk is engaging and the discussion is animated, there is a chance that someone looks out and arrives at the belief that it is rainy. The empirical risk further attenuates the T-practice-based reason to believe that blindspot proposition. What about the strength of E in RAIN? The fact the proposition is a blindspot is a strong evidential reason against believing that proposition, since the belief in that proposition is necessarily false. The normative strength of the T-practice-based reason to believe that blindspot is less than the normative strength of E, and therefore it is neither a requiring nor a permitting reason for Ginger to believe that it is rainy outside and no one in the room believes that it is rainy outside.[23] Similar analysis can be provided to other type (b) cases where the T-practice-based reason to believe a blindspot proposition is significantly attenuated by the fact that it lies at the periphery of the T-practice and is outweighed by the evidential reason against forming a necessarily false belief.

Type (c) cases involve propositions unsupported by evidence. Consider the following example:

> (MOUNTAIN HARE) Fin was hiking in the Cairngorms National Park and saw a large white object that looked like a snowman. It was a snowy day in April and the ground was covered in white. Fin saw that the children of local farmers were out playing in that area. Fin knows that mountain hares have never been seen in this part of the park. Suppose that, luckily for Fin, the large object he spotted was a mountain hare.

In this case, the claim that Fin saw a mountain hare is unsupported by his evidence. The T-practice-based reason for Fin to believe that he saw a mountain hare cannot be a requiring reason, as we have argued in Sect. 6.2.3, since it is incompatible with the evidential norm

---

[23] The normative strength of E is not attenuated by risk attenuators since there is neither empirical nor normative risk that E fails to constitute a reason against believing that blindspot proposition. Thus, E has the same normative strength as that of the T-practice-based reason to refrain from believing that blindspot proposition when it is false from which E derives its normative strength. Since the reason to believe the truth and the reason to avoid falsity are two sides of the same coin, the normative strength of the T-practice-based reason to believe p when p is true should be the same as the normative strength of the T-practice-based reason to refrain from believing p when p is false. In RAIN, since the normative strength of the T-practice-based reason to believe the true blindspot proposition is attenuated while E has the same normative strength as the *unattenuated* T-practice-based reason against believing false propositions, it follows that the normative strength of the T-practice-based reason to believe that blindspot is less than the normative strength of E

which says that one ought to refrain from believing p if p is unsupported by one's evidence and OIC. With variantism, we are able to explain, without the need to invoke those principles, why in MOUNTAIN HARE, the T-practice-based reason does not have a requiring weight.

According to variantism, the empirical risk of the T-practice-based reason is significant given that the evidential probability that Fin saw a mountain hare is low. The risk attenuates the normative strength of the T-practice-based reason for Fin to believe that he saw a mountain hare. What is the normative strength of E in MOUNTAIN HARE? The evidential reason against believing that Fin saw a mountain hare is strong: the large object he spotted looked like a snowman, the children were out playing, and Fin knows that mountain hares have never been seen in this part of the park. The risk of E failing to constitute a reason is low. Putting these considerations together, in MOUNTAIN HARE, the normative strength of the T-practice-based reason is less than the normative strength of E, and therefore it is neither a requiring nor a permitting reason for Fin to believe that he saw a mountain hare. Similar analysis can be provided to other type (c) cases involving propositions unsupported by evidence.

I am moving through all these cases relatively quickly, but the analysis should be clear enough considering what we have said in the previous section. The way in which variantism enables us to explain this wide range cases is an attractive feature of the view: a T-practice-based reason functions holistically just like any other reasons, and its weight varies depending on the presence/absence of conditions, modifiers and the normative strength of opposing evidential reasons. Moreover, it is economical since variantism enables us to explain why the truth of p fails to make a claim on us regarding what we ought to or may believe in type (a)-(c) cases, without having to invoke ought-limiting principles or evidential norms.

Let me briefly raise and respond to an objection to the variantist explanation by appealing to over-intellectualisation before situating my view within the context of similar alternatives in the next section. A critic may suggest that the process of weighting the T-practice-based reason to believe p/refrain from believing not-p when p is true on the variantist account seems overly complicated, and few epistemic agents seem to be capable of engaging in such weighting processes correctly. And yet, most of us seem to be able to judge correctly what sort of claim truth can make on us in a given case. So, something must be wrong with the variantist account.

In response, it is worth emphasizing that the version of variantism I develop and defend in this chapter is simply an application of the widely accepted view that reasons are holistic

and context-sensitive. So if there is indeed a problem of intellectualization, it is a problem for everyone who accepts a holistic account reasons, according to which how weighty a reason is in making a claim on us regarding what to do or believe depends on the circumstances in the way we sketched in Sect.7.1. Moreover, from an evolutionary point of view, it is not surprising that human beings develop capacities to recognize and respond to the relevant conditions, modifiers and opposing reasons in a given situation and to weigh reasons in an efficient and reliable manner. After all, to be able to respond to the claims that reasons make on us is crucial to our survival and flourishing.

Another potential objection comes from proponents of alternative views who claim that their views offer equally good explanation of these cases. In the next section, I will address this worry and strengthen the case for variantism by contrasting it favorably with two alternative views.

7.4 Contrasting Variantism with Alternative Views

7.4.1 Hedged Truth Norms

Some versions of the invariantist view of the truth norm remain unscathed by the arguments discussed in the previous chapter. For example, consider a hedged obligatory account of the truth norm which says that the truth of p has a requiring weight *except* (1) when S cannot believe p or refrain from believing not-p; (2) when p is a blindspot proposition; and (3) when p is not supported by evidence available to S. The hedged obligatory account of the truth norm vindicates the intuitive verdict that in TIME PASSING, RAIN and MOUNTAIN HARE the truth of the proposition involved does not have a requiring weight: these are cases to which the obligatory account does not apply.

One problem with the hedged obligatory account of the truth norm is that it faces further counterexamples. Consider the following case:

> (MOUNTAIN HARE*) Fin was hiking in the Cairngorms National Park and saw a large white object that looked like a mountain hare. It was a snowy day in April and the ground was covered in white. Fin believes that he saw a mountain hare. However, Fin knows that mountain hares have never been seen in this part of the park. Furthermore, it is unusual for a mountain hare to have white fur at this time of year. Luckily for Fin, the large object he spotted was indeed a mountain hare. Based on that belief Fin conducted further scientific research in that area and made new findings about mountain hares.

In this case, it seems that the T-practice-based reason for Fin to believe that he saw a mountain hare is a permitting but not a requiring reason. He does not seem to commit any epistemic wrongdoing if he did not form the believe that he saw a mountain hare, after all, there is evidence suggesting that he did not see a mountain hare. But according to the hedged obligatory account of the truth norm, the T-practice-based reason is a requiring reason. We can conceive of other cases in which the truth of a proposition falls short of having a requiring weight but are not included in the list of exceptions like (1)–(3).

Now, you might think this problem can be solved by adding further exceptions to the hedged principle. It is designed to include all potential exceptions so that it can explain away any putative counterexample. But there are good reasons to prefer variantism over a hedged view with a long list of exceptions.

To begin with, variantism is explanatorily more powerful. Hedged obligatory account of the truth norm does not have the capacity to offer an analysis of the weight of truth in type (a)-(c) cases, for it does not apply in those cases. By contrast, variantism enables us to say what sort of claim truth can make on us regarding to what to believe in those cases, and why.

There is also a concern about whether non-trivial principles can include exceptions. On the one hand, one might think it is not possible to draw a complete list of exceptions. And on the other hand, if the list of exceptions is open-ended, there is a further worry that a hedged principle including an opened list of exceptions seems incapable of doing any explanatory work.[24]

More generally, hedged invariantist truth norms lack motivations on a holistic framework. On a holistic framework, no reason has pre-fixed weight outside a given context – the weight of a reason is neither prior nor independent of other salient facts which obtain in the context. Whether a reason has the weight it has depends on what else is true in the situation. The weight of the T-practice-based reason is neither determined nor explained by some invariantist norms, but depends on the relevant conditions, modifiers and its opposing reasons in the way I spelled out in the previous section. An invariantist view on a holistic framework does no real explanatory work.

Furthermore, on a holistic framework, the invariantist has the burden to show that the T-practice-based reason has the same requiring/permitting weight in all contexts, despite the

[24] This move is structurally parallel to that adopted by some ethicists in defence of invariantist moral principles. See, McKeever & Ridge (2008).

infinitely many possible changes of circumstances. It calls out for explanation why such incredible coincidence should obtain. Variantism, by contrast, is simply the result of applying the holism to T-practice-based reasons for belief. So, if one accepts the holistic account of reason, as many do, it is difficult to see why one would prefer a hedged invariantist truth norm over variantism.

One might contend that even if an invariantist truth norm on a holistic framework does no explanatory work, it nevertheless plays an important role in practice by guiding our deliberation. For example, you might think that the hedged obligatory account of the truth norm, if true, tells us under what circumstances, i.e., when a proposition is true and is not a blindspot proposition and so on, we ought to believe that proposition. An invariantist truth norm (bracketing all putative counterexamples) seems the to be just the kind of principle we need to consult in making up our mind about what to believe. However, in part II of the thesis, we have challenged this inferentialist picture of guidance, according to which, to be guided by a norm is to make an inference in accordance with that norm. We have argued that for an agent to be guided by the truth norm, it is not necessary for that agent to invoke the norm in her doxastic deliberation. She can be guided by the truth norm if she responds to the T-practice-based reason, which is to manifest the relevant reason-responsive dispositions in a situation or to manifest the distinctive phenomenology of believing something for the T-practice-based reason. There is no motivation for a hedged invariantist truth norm from the practical point of view either.

7.4.2 Perspectivism

Recently, an increasingly popular view in the literature of normativity is perspectivism. Perspectivists hold that what one ought to do or believe depends on one's epistemic perspective. Objectivists, by contrast, hold that what one ought to do or believe depends on all the facts, irrespective of one's epistemic perspective.[25] There are, of course, many forms of perspectivism. What interests me here is a form of perspectivism about the truth norm, which, as we will see, shares a number of features with variantism.

Perspectivism about the truth norm is the view that, roughly, whether the truth of a proposition is a requiring (or permitting) reason for one to believe that proposition is

[25] Perspectivism is motivated by two main arguments: one from counterexamples to objectivism and the other from guidance. For further discussion of these arguments, see, e.g., Jackson (1991), Kiesewetter (2011, 2018), McHugh & Way (2017), Way & Whiting (2016, 2017), Littlejohn (2019).

determined by one's epistemic perspective, which is constituted by the total evidence available to one in a given situation (e.g., Conee & Feldman 2004; Gibbons 2013; McHugh & Way 2017).[26] It is easy to see the overlap between variantism and perspectivism. Both variantists and perspectivists reject unqualified invariantist truth norms such as the obligatory and the permissible account of the truth norm. Both hold that the correct truth norm is context-sensitive. And both think that one's epistemic perspective matters to what one ought to or may believe in a given situation. So, it will help to further highlight the distinct virtue of variantism by comparing it with perspectivism about the truth norm, which is in its close vicinity.

To facilitate our discussion, it will be helpful to put a more precise formulation of perspectivism on the table. We can focus on the following version of perspectivism:

> (Perspectivism) For any S, p, the truth of p is a requiring reason for S to believe p if and only if S's evidence decisively supports p; and the truth of p is a permitting reason for S to believe p if and only if S's evidence sufficiently supports p.[27]

As one might expect, perspectivism does well with respect to type (c) cases, such as MOUNTAIN HARE. In MOUNTAIN HARE, Fin's evidence does not sufficiently support the claim that he saw a mountain hare. So according to perspectivism, the truth that Fin saw a mountain hare is neither a requiring nor a permitting reason. It can also explain why in MOUNTAIN HARE*, the truth that Fin saw a mountain hare is a permitting reason for him to believe that. Fin saw a white object that looked like a mountain hare. But since Fin knows that mountain hares have never been seen in the area, Fin's evidence seems to sufficiently but not decisively support the claim that he saw a mountain hare. So according to perspectivism, the T-practice-based reason has a permitting but not a requiring weight.

It is less clear, however, that perspectivism can explain all type (a) cases involving propositions which one cannot believe/refrain from believing. A subset of type (a) cases involving true propositions decisively/sufficiently supported by one's evidence provides counterexamples to perspectivism, because according to perspectivism, the T-practice-based

---

[26] Alternative evidence-based construals of epistemic perspective include what one knows, what one is in a position to know, what one justifiably believes and so on. The difference does not matter for our present purposes.

[27] This, of course, raises questions about what it takes to provide *decisive* and *sufficient* evidential support for a proposition. But it is not my goal to develop and defend perspectivism here. Our intuitive understanding of these notions will suffice for our purposes.

reasons to believe p in those cases has a requiring/permitting weight. But we have seen, when one cannot believe p, the T-practice-based reason to believe p is disabled and thus does not have a requiring/permitting weight.

Perspectivism also has difficulty explaining all type (b) cases involving blindspot propositions. A subset of type (b) cases involving true blindspot propositions decisively/sufficiently supported by one's evidence also provides counterexamples to perspectivism, because according to perspectivism, those true propositions have a requiring/permitting weight. But we have seen, when p is a blindspot proposition, the T-practice-based reason is significantly attenuated and outweighed by opposing reasons against believing p. Thus, it does not have a requiring/permitting weight.

Perspectivists might respond to the putative counterexamples by tightening up their account. For example, consider the following revised version of perspectivism:

> (Perspectivism*) For any S, p, the truth of p is a requiring reason for S to believe p if and only if the evidence *possessed* by S decisively supports p; and the truth of p is a permitting reason for S to believe p if and only if the evidence *possessed* by S sufficiently supports p.

According to one recent proposal of evidence *possession*, epistemic access to evidence is insufficient for one to count as possessing that evidence. An additional, dispositionalist condition must be satisfied. If one possesses evidence for p, then one must treat that evidence as a reason for believing p, which is to manifest a disposition to believe p when that evidence constitutes a reason for believing p (e.g., Sylvan 2016a; Lord 2018). So, on this account, the truth of p is a requiring/permitting reason for S to believe p if and only if S treats the decisive/sufficient evidence available to S as a reason for believing p.

Now, perspectivists could argue that perspectivism* can explain all type (a) and type (b) cases. Consider type (a) cases involving propositions which one cannot believe. Perspectivists could argue that, even if one has epistemic access to sufficient/decisive evidence for a proposition in those cases, it is impossible for one to treat the evidence as a reason for believing that proposition since one cannot have the relevant disposition to believe something that is impossible for one to believe. Hence, it is impossible for one to possess such evidence. So according to perspectivism*, the T-practice-based reasons to believe p in those type (a) cases does not have a requiring/permitting weight.

Considering type (b) cases involving blindspot propositions, perspectivists could argue that, even if one has epistemic access to sufficient/decisive evidence for a proposition in

those cases, it is impossible for one to treat that evidence as a reason for believing a blindspot proposition since one cannot have the relevant disposition to believe something that, one knows, if she were to believe it, would be necessarily false. Hence, it is impossible for one to possess such evidence. So according to perspectivism*, the T-practice-based reason to believe p in those type (b) cases does not have a requiring/permitting weight.

So, it seems that perspectivism* can do just as well as variantism in explaining all type (a)-(c) cases. Why then, should we prefer variantism? We cannot properly assess the dispositionalist account of evidence possession here, which is the key to the success of the revised perspectivist explanation of those cases. But even if perspectivism* can successfully explain type (a)-(c) cases, there remains a crucial difference between variantism and perspectivism*. On variantism, as I have developed it, non-evidential considerations play important roles in weighting the T-practice-based reason to believe p/refrain from believing not-p when p is true (i.e., as enabling/disabling conditions and as attenuating/intensifying modifiers), whereas on perspectivism*, only evidence can play such a role.

One important virtue of variantism, then, is that it can explain two further type of cases. Consider first type (d) cases involving true propositions that matter to us but are not well-supported by the evidence possessed by the agent. Variantism vindicates our intuitive judgment that the truth of the matter can make a strong claim on us regarding what to believe in type (d) cases. As we have argued in Sect. 7.3, when the survival and flourishing of our community hinges on us grasping the truth about p, even if p may be painful to accept or hard to uncover, the truth of p has a requiring weight. For example, the fact that p concerns issues such as the Covid-19 pandemic, climate change, social justice, etc., significantly intensifies the T-practice-based reason to believe p/refrain from believing not-p when p is true. In such cases, believing p and refraining from believing not-p when p is true is crucial to promote social cooperation, knowledge production and maintenance. Thus, according to variantism, in type (d) cases, the normative strength of T-practice-based reason outweighs the normative strength of its opposing evidential reasons and therefore has a requiring weight. Perspectivism* by contrast cannot explain type (d) cases. According to perspectivism*, the T-practice-based reason makes no claim on what one ought to believe in those cases, since one does not possess the relevant evidence for those propositions.

Variantism can also explain type (e) cases involving true propositions that matter very little to us but are well-supported by the evidence possessed by the epistemic agent. Variantism vindicates our intuitive judgment that in such cases, the truth of the matter is unlikely to make a requiring claim on us regarding what to believe. For example, the fact that

p concerns issues such as car numbers significantly attenuates the T-practice-based reason to believe a proposition about car numbers. Believing the truth about trivial issues such as car numbers is close to the periphery of the T-practice since it plays limited role in promoting social cooperation, knowledge production and maintenance. Thus, according to variantism, in type (e) cases, the normative strength of T-practice-based reason is unlikely to outweigh the normative strength of its opposing reasons and therefore it is unlikely to have a requiring weight. Perspectivism*, by contrast, cannot explain our intuition about type (e) cases.

Thus, even if perspectivism* about the truth norm shares a number of features with variantism and can deal with type (a)-(c) cases, there remains a crucial difference. Variantism has a distinct virtue in accommodating type (d) and (e) cases. Variantism therefore offers a more attractive response to the weighting challenge than hedged truth norms and the two versions of perspectivism we considered.

7.5 Conclusion

I began this chapter by suggesting that a natural way to spell out variantism is to tap into the widely accepted view in metaethics, which says that reasons are holistic. On this view, whether a reason has a requiring or permitting weight depends on the circumstances. Variantism is simply the result of applying the view to the T-practice-based reason. Truth as a reason for belief functions holistically just like any other reasons: its weight in making a claim on us regarding what to believe can vary depending on the presence/absence of conditions, modifiers and opposing reasons in a given situation. I then spelled out in detail how the T-practice-based reason can be conditioned, modified and weighted by opposing evidential reasons. I put variantism to work and argued that variantism enables us to explain the three kinds of cases which make trouble for the invariantist accounts of the truth in an attractive and economical way. Finally, I contrasted variantism favourably with two alternative views and highlighted the distinct virtue of variantism. Variantism therefore offers an attractive response to the weighting challenge.

8. Summary of the Thesis and Future Work

The main goal of this thesis has been to address some recent challenges relating to various aspects of the truth norm of belief on a reason-based normative framework. In doing so, I developed and defended a practice-based, variantist account of the truth norm, according to which, the truth norm of belief is grounded in a justified social practice, guides our belief-formation on a reason-responsive model of epistemic guidance, and makes claims on us regarding what we ought to/may believe depending on the circumstances.

Chapter 1 focused on setting the stage. I began by introducing three recent anti-normativist challenges to the truth norm of belief concerning its authority, guiding capacity and normative force. I explained what I take these challenges to involve on a reason-based normative framework. On a reason-based normative framework, normativity is a matter of (normative) reasons. Reasons are facts that count in favour of a response, grounded in value-based, desire-based, constitutive, or practice-based facts. Reasons are the sort of things that we can respond to and they have weights in a making a claim on us when we stand in a relation to them. Three anti-normativist challenges arise in recent discussion. First, the grounding challenge concerns whether the reason to believe truly (correctly) and to refrain from believing falsely (incorrectly) can be properly grounded. Second, the guidance challenge concerns whether the reason to believe truly and to refrain from believing falsely is something we can respond to in our belief-formation. Third, the weighting challenge concerns the weight of the reason to believe truly and to refrain from believing falsely in making a claim on us regarding what we ought to/may believe.

In Part I, I addressed the grounding challenge. In Chapter 2, I discussed and criticised two main views about the authority of the truth norm in the current literature: the evaluative and the constitutive account of the truth norm. According to the evaluative account, the truth norm of belief is authoritative because true beliefs are good. It attempts to ground the authority of the truth norm by appealing to value-based/desire-based reasons to believe truly and only truly. This view faces the wrong-kind-of-reason problem and commits itself to a substantial view about the nature of truth, i.e., that truth is something valuable/desirable. According to the constitutive account, the truth norm of belief is authoritative because it is *constitutive* of belief that belief is subject to a norm of truth. It attempts to ground the authority of the truth norm by appealing to constitutive reasons to believe truly and only truly. This view faces the problem of limited 'authority' and it commits itself to a substantial view about the nature of belief, i.e., that belief is a type of mental attitude that is essentially subject to a

norm of truth, in combination with some version of representationalist, functionalist, or dispositionalist account of belief. I identified and explained what I take to be a more fundamental concern for both views, which has not received much discussion. On the reason-based framework, the fundamental challenge for both the evaluative and the constitutive account is to show that the relevant reason grounding facts obtain. I rejected three arguments for the claim that the reason to believe truly and to refrain from believing falsely is value-based/desired-based and three arguments for the claim that the reason is constitutive. I concluded that proponents of the two existing views fail to provide satisfactory answers to the grounding challenge for they are yet to convince us that their preferred reason grounding facts obtain.

In Chapter 3, I argued for a practice-based view about the authority of the truth norm. On my view, the authority of the truth norm is grounded in a justified social practice, i.e., the T-practice. The practice-based account is preferable to the evaluative and the constitutive account. It avoids the wrong-kind-of-reason problem and the problem of limited 'authority'. It is a view that reflects the social aspect of belief and aligns with the growing trend of social epistemology – an approach to epistemology which focuses on the social dimension of knowledge (belief) acquisition and transmission. The reason to believe truly and only truly is practice-based. Along the way, I argued that a social practice has three earmarks: it is rule-governed, interactive and has social functions. I introduced the method of genealogy for uncovering the function of a social practice and applied it to our analysis of the T-practice. I showed that the T-practice functions to facilitate knowledge production, maintenance, and social production, which is indispensable to the survival and flourishing of human societies. Using Orwell's *1984* as a thought experiment, I argued that even a partial breakdown of the T-practice would bring substantial harm to its participants. The T-practice, despite not being flawless, is a justified social practice and grounds the authority of the truth norm of belief.

In Part II, I addressed the guidance challenge. Chapter 4 began with a discussion of the No Guidance argument. I briefly considered a response to the argument which maintains the inferentialist model of epistemic guidance. On the inferentialist model, to be guided by the truth norm to form a belief is to infer that belief in accordance with the truth norm. I identified and explained what I take to be a significant concern for the inferentialist model of epistemic guidance, which has not received much discussion. I argued that the inferentialist model faces difficulties in light of emerging empirical evidence about our system of belief-formation. To offer an alternative response to the guidance challenge, I proposed a reason-responsive model of guidance, according to which to be guided by an epistemic

norm to form a belief is to respond to the relevant normative reasons. I argued that the reason-responsive model provides an attractive alternative to the inferentialist model – it meets all four criteria for a viable model of epistemic guidance.

Chapter 5 tackled the Novice Problem and defended the idea that the truth norm of belief can provide guidance on the reason-responsive model of guidance. The problem arises because a novice of the T-practice has not yet possessed the relevant dispositions to believe in response to the T-practice-based reason by the very definition of a novice, so she cannot be said to be guided by the truth norm on the account I gave in Chapter 4. I provided a solution to the Novice Problem. In doing so, I developed a novel hybrid dispositionalist/phenomenological account of what it takes to respond to the T-practice-based reason. My central idea is that believing p in response to the T-practice-based reason is a mental action. To respond to the T-practice-based is to either manifest the relevant disposition to respond to the T-practice-based reason or manifest a distinctive phenomenology – an action-awareness of believing p which is distinctively non-observational. The non-observational action-awareness has two features: (1) it is as of something that has a certain direction – moving toward the aim of believing something true and (2) it is as of something that is subject to some self-imposed constraints in the service of achieving that aim. The hybrid dispositionalist/phenomenological account provides a solution to the Novice Problem while retaining the virtues a purely dispositionalist reason-responsive model enjoys over the inferentialist model of epistemic guidance.

In Part III, I addressed the weighting challenge. In Chapter 6, I discussed two leading accounts about the normative force of the truth norm in the current literature: the obligatory and the permissible account of the truth norm. According to the obligatory account of the truth norm, the T-practice-based reason to believe truly and only truly is a requiring reason. The obligatory account has been criticized on the grounds that it makes demands on us that are excessive, impossible to be satisfied and clash with our epistemic obligation to believe in accordance with evidence. These problems have been thought to favour the permissible account of the truth norm, according to which the T-practice-based reason to believe truly and only truly is a permitting reason. I argued that the permissible account faces analogous problems and does not fare better than the obligatory account that it aims to replace. I offered a diagnosis: on the reason-based normative framework, both views are invariantist, for they claim that the T-practice-based reason has a certain weight across all circumstances. But the T-practice-based does not always have the same requiring/permitting weight across all circumstances – the weight of the T-practice-based reason may vary from case to case. I

considered and rejected two potential responses invariantists may make in response. One is to appeal to alternative conceptions of doxastic oughts that are unconstrained by ought-limiting principles. Another is to reject the principle of 'ought' implies 'can'. I concluded that there is no easy fix of the problems facing the invariantist.

Chapter 7 developed and defended a variantist account of the truth norm which says that whether the T-practice-based reason is a requiring or permitting reason for belief can vary depending on the circumstances. A natural way to spell out variantism is to tap into the widely accepted view in metaethics, which says that reasons are holistic. Truth as a reason for belief functions holistically just like any other reasons: its weight in making a claim on us regarding what to believe can vary depending on the presence/absence of conditions, modifiers and opposing reasons in each situation. I spelled out in detail how the T-practice-based reason can be conditioned, modified and weighted by opposing evidential reasons. I put variantism to work and argued that variantism enables us to explain the three kinds of cases which make trouble for the invariantist accounts of the truth in an attractive and economical way. Variantism is also preferable to hedged truth norms and perspectivism. I concluded that variantism offers the most attractive response to the weighting challenge.

I believe this practice-based, variantist account of the truth norm of belief on a reason-based framework advances our understanding of the normative relation between belief and truth, and it opens several avenues for future research. I will conclude the thesis by briefly sketching some of them.

Although our discussion does not imply that the normative relation between belief and truth is *not* internal to either belief or truth, it strongly suggests so. The normative relation obtains not because of the intrinsic nature of belief or truth but because believing truly and refraining from believing falsely constitutes a justified social practice, which emerged and evolved in response to our epistemic needs. While my focus has been on developing a practice-based account of the truth norm of belief, the practice-based approach seems promising in relation to other epistemic norms. It invites further questions such as whether evidential norms can be grounded in the T-practice or whether they are grounded in a social practice that is closely related to yet independent from the T-practice. Is the truth norm of belief the most fundamental norm of belief? And if so, can we explain why this is the case on the reason-based normative framework?

In Part II, I developed the idea that believing something in response to the T-practice-based reason is a kind of mental action and suggested that its distinctive phenomenology can manifest the agent's reason-responsiveness. This helped us to explain how a novice may be

guided by the truth norm when she has yet to acquire the relevant reason-responsive dispositions. While my focus has been on developing a dispositionalist/phenomenological account of responding to the T-practice-based reason, the basic idea may be generalized. One project for future research is to investigate whether φ-ing in response to reasons in general has a distinctive phenomenology and what it may involve. One may also raise questions about the relationship between the dispositionalist and the phenomenological condition. Is one condition prior to the other condition? And are there cases where one condition is relevant while the other is not in determining whether one's belief is reason-responsive? There is also a further complication regarding apparent epistemic reasons – facts that seem to count in favour of believing p but are not. Is someone who responds to an apparent T-practice-based reason in forming a belief guided by the truth norm? These questions have to be left for future research.

Considering our discussion in Part III, one issue that I have not dealt with is how to calibrate the account of guidance around the varying weight of the T-practice-based reason. There is a further question about what it takes to respond to a reason with a certain weight in each situation. It seems that to respond to a reason with a certain weight involves evaluation of the various conditions, modifiers and opposing reasons in the situation under consideration. For example, one type of modifiers I discussed is risk attenuator. A reason for belief may be empirically risky or/and normatively risky. The higher the risk, the lower the weight of that reason in making a claim on us regarding what we ought to/may do. So, one way to begin thinking about what it takes to respond to a reason with a certain weight is to think about how we should respond to risk.

No doubt, our discussion invites more questions than the few I sketched above. But I believe this thesis has made some steps towards a better understanding of the truth norm of belief. The reason-based normative framework allows us to see what underlies the issues about authority, guidance and normative force and to tackle them head-on. Truth is a genuine norm of belief – it is authoritative, guiding, and makes claims on us regarding what we ought to or may believe.

Bibliography

Adler, J. 2002. *Belief's Own Ethics*. Cambridge, MA: MIT Press.

Ahkhmatova, Anna. 1988. *Anna Akhmatova Selected Poems*. Translated by D.M. Thomas. London: Peguine Books.

Alicke, Mark D., and Constantine Sedikides. 2009. "Self-enhancement and self-protection: What they are and what they do." *European Review of Social Psychology* 20 (1): 1-48.

Alston, William P. 1985. "Concepts of Epistemic Justification." *The Monist* 68 (2): 57-89.

---. 1988. "The deontological conception of epistemic justification." *Philosophical Perspectives* 2 (n/a): 257-299.

Alvarez, Maria. 2018. "Reasons for action, acting for reasons, and rationality." *Synthese* 195 (8): 3293-3310.

Anscombe, G. E. M. 1957. *Intention*. Cambridge, MA: Harvard University Press.

Armstrong, D. M. 1968. *A Materialist Theory of the Mind*. London: Routledge.

---. 1973. *Belief, Truth and Knowledge*. Cambridge Cambridge University Press.

---. 2004. *Truth and Truthmakers*. Cambridge: Cambridge University Press.

Arpaly, N., and T. Schroeder. 2012. "Deliberation and Acting for Reasons." *Philosophical Review* 121 (2): 209-239.

Arpaly, Nomy, and Timothy Schroeder. 2013. *In Praise of Desire*. Oxford: Oxford University Press.

---. 2015. "A Casual Theory of Acting for Reasons." *American Philosophical Quarterly* 52 (2): 103-114.

Audi, Robert. 1993. *The Structure of Justification*. Cambridge: Cambridge University Press.

Bader, Ralf. 2016. "Conditions, Modifiers, and Holism." In *Weighing Reasons*, edited by Errol Lord and Barry Maguire. Oxford: Oxford University Press.

Barnes, Annette. 1997. *Seeing Through Self-Deception*. Cambridge: Cambridge University Press.

Barry, Melissa. 2007. "Realism, Rational Action, and the Humean Theory of Motivation." *Ethical Theory and Moral Practice* 10 (3): 231-242.

Basu, Rima. 2018. "Can Beliefs Wrong?" *Philosophical Topics* 46 (1): 1-17.

---. 2019. "What We Epistemically Owe To Each Other." *Philosophical Studies* 176 (4): 915–931.

Beebee, Helen, and Julian Dodd. 2005. *Truthmakers: The Contemporary Debate*. Oxford: Clarendon Press.

Bermúdez, José Luis. 2000. "Self-deception, Intentions and Contradictory Beliefs." *Analysis* 60 (4): 309-319.

Boghossian, Paul A. 2003. "The Normativity of Content." *Philosophical Issues* 13 (1): 31-45.

---. 2008. "Epistemic Rules." *Journal of Philosophy* 105 (9): 472-500.

Boghossian, Paul, and Crispin Wright. ms. "The Adoption Problem."

BonJour, Laurence. 1985. *The Structure of Empirical Knowledge*. Cambridge, MA: Harvard University Press.

Bortolotti, Lisa. 2009. *Delusions and Other Irrational Beliefs*. Oxford: Oxford University Press.

Boyle, Matthew. 2011. "'Making up Your Mind' and the Activity of Reason." *Philosophers' Imprint* 11 (17): 1-24.

Brandom, Robert B. 1994. *Making It Explicit: Reasoning, Representing, and Discursive Commitment*. Cambridge, MA: Harvard University Press.

Buckwalter, Wesley, and John Turri. 2015. "Inability and Obligation in Moral Judgment." *PLoS ONE* 10 (8): 1-20.

Burge, Tyler. 2010. *Origins of Objectivity*. Oxford: Oxford University Press.

Bykvist, Krister. 2017. "Moral Uncertainty." *Philosophy Compass* 12 (e12408): 1-8.

Bykvist, Krister, and Anandi Hattiangadi. 2007. "Does Thought Imply Ought?" *Analysis* 67 (296): 277–285.

---. 2013. "Belief, Truth, and Blindspots " In *The Aim of Belief*, edited by Timothy Chan, 100-122. Oxford: Oxford University Press.

Carruthers, Peter. 2000. *Phenomenal Consciousness: A Naturalistic Theory*. Cambridge: Cambridge University Press.

---. 2011. *The Opacity of Mind: An Integrative Theory of Self-Knowledge*. Oxford: Oxford University Press.

Chan, David K. 1995. "Non-Intentional Actions." *American Philosophical Quarterly* 32 (2): 139 - 151.

Chituc, Vladimir, Paul Henne, Walter Sinnott-Armstrong, and Felipe De Brigard. 2016. "Blame, Not Ability, Impacts Moral "Ought" Judgments for Impossible Actions: Toward an Empirical Refutation of "Ought" Implies "Can"." *Cognition* 150: 20-25.

Chrisman, Matthew. 2008. "Ought to Believe." *Journal of Philosophy* 105 (7): 346-370.

Christensen, David. 2008. "Does Murphy's Law Apply in Epistemology." *Oxford Studies in Epistemology* 2: 3-31.

Conee, Earl. 1992. "The Truth Connection." *Philosophy and Phenomenological Research* 52 (3): 657-669.

Conee, Earl, and Richard Feldman. 2004. *Evidentialism: Essays in Epistemology*. Oxford: Oxford University Press.

Copp, David. 1997. "Defending the Principle of Alternate Possibilities: Blameworthiness and Moral Responsibility." *Noûs* 31 (4): 441-456.

---. 2008. "'Ought' Implies 'Can' and the Derivation of the Principle of Alternate Possibilities." *Analysis* 68 (297): 67–75.

Craig, Edward. 1990. *Knowledge and the State of Nature: An Essay in Conceptual Synthesis*. Oxford: Oxford University Press.

Crisp, Roger. 2000. "Particularizing Particularism." In *Moral Particularism*, edited by Brad Hooker and Margaret Olivia Little, 23-47. Oxford: Oxford University Press.

Dancy, Jonathan. 2004. *Ethics Without Principles*. Oxford: Oxford University Press.

Darwall, Stephen L. 1983. *Impartial Reason*. Ithaca, NY: Cornell University Press.

Davidson, Donald. 1970. *Essays on Actions and Events*. Oxford: Clarendon Press.

---. 1984. *Inquiries Into Truth And Interpretation*. Oxford: Oxford University Press.

---. 1985. "Incoherence and Irrationality." *Dialectica* 39 (4): 345-54.

DePaul, M. 2009. "Does an Ugly Analysis Entail that the Target of the Analysis Lacks Value?" In *Epistemic Value*, edited by Pritchard, Haddock and Millar, 112-38. Oxford: Oxford University Press.

Douglas, Karen M., Joseph E. Uscinski, Robbie M. Sutton, Aleksandra Cichocka, Turkay Nefes, Chee Siang Ang, and Farzin Deravi. 2019. "Understanding Conspiracy Theories." *Political Psychology* 40 (S1): 3-35.

Dretske, Fred. 1988. *Explaining Behavior: Reasons in a World of Causes*. Cambridge, MA: MIT Press.

Dummett, Michael. 1959. "Truth." *Proceedings of the Aristotelian Society* 59 (1): 141-62.

Ebert, Philip A., and Martin Smith. 2012. "Introduction: Outright Belief and Degrees of Belief." *Dialectica* 66 (3): 305-308.

Ebert, Philip A., Martin Smith, and Ian Durbach. 2020. "Varieties of Risk." *Philosophy and Phenomenological Research* 101 (2): 432-455.

Egan, Andy. 2008. "Seeing and Believing: Perception, Belief Formation and the Divided Mind." *Philosophical Studies* 140 (1): 47 - 63.

Engel, Pascal. 2002. *Truth*. Montreal Mcgill-Queen's University Press.

---. 2013. "In Defence of Normativism about the Aim of Belief." In *The Aim of Belief*, edited by Timothy Chan, 32-63. Oxford: Oxford University Press.

Evans, Jonathan St. B. T., and Keith E. Stanovich. 2013. "Dual-Process Theories of Higher Cognition:Advancing the Debate." *Perspectives on Psychological Science* 8 (3): 223-241.

Fassio, Davide. 2011. "Belief, Correctness and normativity." *Logique Et Analyse* 54 (216): 471-486.

Feldman, Richard. 1988. "Epistemic Obligations." *Philosophical Perspectives* 2 (n/a): 235-256.

---. 2000. "The Ethics of Belief." *Philosophy and Phenomenological Research* 60 (3): 667-695.

Feldman, Richard, and Earl Conee. 1985. "Evidentialism." *Philosophical Studies* 48 (1): 15 - 34.

Field, Hartry. 1989. *Realism, Mathematics & Modality*. Oxford: Blackwell.

Finlay, Stephen. 2006. "The Reasons that Matter." *Australasian Journal of Philosophy* 84 (1): 1 – 20.

Fodor, Jerry A. 1968. *Psychological Explanation: An Introduction To The Philosophy Of Psychology*. New York, NY: Random House.

---. 1987. *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge, MA: MIT Press.

---. 1990. *A Theory of Content and Other Essays*. Cambridge, MA: MIT Press.

Fogal, Daniel. 2016. "Reasons, Reason, and Context." In *Weighing Reasons*, edited by Errol Lord and Barry Maguire. Oxford: Oxford University Press.

Frankfurt, Harry G. 1969. "Alternate Possibilities and Moral Responsibility." *Journal of Philosophy* 66 (23): 829.

Frege, Gottlob. 1948. "Sense and Reference." *Philosophical Review* 57 (3): 209-230.

---. 1956. "The Thought: A Logical Inquiry." *Mind* 65 (259): 289-311.

---. 1991. *Posthumous Writings*. Oxford: Wiley-Blackwell.

---. 2013. *Basic Laws of Arithmetic*. Translated by Philip A. Ebert. Oxford: Oxford University Press.

Fricker, Miranda. 2007. *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford: Oxford University Press.

Fricker, Miranda, Peter Graham, David Henderson, and Nikolaj Jang Pedersen. 2019. *The Routledge Handbook of Social Epistemology*. London: Routledge.

Gert, Joshua. 2003. "Requiring and justifying: Two dimensions of normative strength." *Erkenntnis* 59 (1): 5 - 36.

---. 2007. "Normative Strength and the Balance of Reasons." *Philosophical Review* 116 (4): 533-562.

Gertler, Brie. 2001. "Introspecting Phenomenal States." *Philosophy and Phenomenological Research* 63 (2): 305-28.

Gibbard, Allan. 1990. *Wise Choices, Apt Feelings: A Theory of Normative Judgment*. Cambridge, MA: Harvard University Press.

---. 2003. "Thoughts and Norms." *Philosophical Issues* 13 (1): 83-98.

---. 2005. "Truth and Correct Belief." *Philosophical Issues* 15 (1): 338–350.

Gibbons, John. 2010. "Seeing What You're Doing." In *Oxford Studies in Epistemology*, edited by T. Szabo Gendler and J. Hawthorne. Oxford: Oxford University Press.

---. 2013. *The Norm of Belief*. Oxford University Press.

Gilbert, Daniel. 1991. "How mental systems believe." *American Psychologist* 46: 101-119.

Gilbert, Daniel, Romin Tafarodi, and Patrick Malone. 1993. "You can't not believe everything you read." *Journal of Personality and Social Psychology* 65 (2): 221-233.

Gilovich, Thomas, Dale Griffin, and Daniel Kahneman. 2002. *Heuristics and Biases: The Psychology of Intuitive Judgment*. Cambridge: Cambridge University Press.

Glüer, Kathrin, and Åsa Wikforss. 2009. "Against Content Normativity." *Mind* 118 (469): 31-70.

---. 2010. "The Truth Norm and Guidance: a Reply to Steglich-Petersen: Discussions." *Mind* 119 (475): 757-761.

---. 2013. "Against Belief Normativity." In *The Aim of Belief*, edited by Timothy Chan, 80-99. Oxford: Oxford University Press.

Goldman, Alvin. 1979. "What is Justified Belief." In *Justification and Knowledge*, edited by George Pappas, 1-23. Springer, Dordrecht.

---. 2008. "Immediate Justification and Process Reliabilism." In *Epistemology: New Essays*, edited by Quentin Smith, 63-82. Oxford: Oxford University Press.

Goldman, Alvin I. 1999. *Knowledge in a Social World*. Oxford: Oxford University Press.

Goldman, Alvin, and Dennis Whitcomb. 2010. *Social Epistemology: Essential Readings*. Oxford: Oxford University Press.

Gowans, Christopher. 1987. "Introduction. The Debate on Moral Dilemmas." In *Moral Dilemmas*, edited by Christopher W. Gowans, 3-33. Oxford: Oxford Uiversity Press.

Graham, P. A. 2011. "'Ought' and Ability." *Philosophical Review* 120 (3): 337-382.

Greenberg, Alexander. 2020. "Should I Believe All the Truths?" *Synthese* 197 (8): 3279-3303.

Haack, Susan. 1976. "The Pragmatist Theory of Truth." *British Journal for the Philosophy of Science* 27 (3): 231-249.

Harman, Gilbert. 1986. *Change in View: Principles of Reasoning*. Cambridge, MA: MIT Press.

Haslanger, Sally. 2018. "What is a Social Practice?" *Royal Institute of Philosophy Supplement* 82: 231-247.

Hasson, Uri, Joseph P. Simmons, and Alexander Todorov. 2005. "Believe It or Not:On the Possibility of Suspending Belief." *Psychological Science* 16 (7): 566-571.

Hattiangadi, Anandi. 2010. "The Love of Truth." *Studies in History and Philosophy of Science Part A* 41 (4): 422-432.

Hazlett, Allan. 2013. *A Luxury of the Understanding: On the Value of True Belief*. Oxford: Oxford University Press.

Helton, Grace. 2020. "If You Can't Change What You Believe, You Don't Believe It." *Noûs* 54 (3): 501-526.

Henne, Paul, Vladimir Chituc, Felipe De Brigard, and Walter Sinnott-Armstrong. 2016. "An Empirical Refutation of 'Ought' Implies 'Can'." *Analysis* 76 (3): 283-290.

Hieronymi, Pamela. 2005. "The Wrong Kind of Reason." *Journal of Philosophy* 102 (9): 437 - 457.

Hobbes, Thomas. 1994. *Leviathan*. edited by E. Curley. Indianapolis: Hackett.

Horwich, Paul. 1998. *Truth*. Oxford: Clarendon Press.

Hume, David 2000. *A Treatise of Human Nature*. edited by David Fate Norton and Mary J. Norton. Oxford: Oxford University Press.

Hursthouse, Rosalind. 1999. *On Virtue Ethics*. Oxford: Oxford University Press.

Jackson, Frank. 1991. "Decision-theoretic Consequentialism and the Nearest and Dearest Objection." *Ethics* 101 (3): 461-482.

James, William. 1907. "Pragmatism's Conception of Truth." *Journal of Philosophy, Psychology and Scientific Methods* 4 (6): 141-155.

---. 2007. *The Principles of Psychology*. Dover Publications. 1890.

Joyce, Richard. 2006. *The Evolution of Morality*. Cambridge, MA: MIT Press.

Kahneman, Daniel, and Shane Frederick. 2002. "Representativeness revisited: Attribute substitution in intuitive judgment." 49-81. Cambridge: Cambridge University Press.

Kant, Immanuel. 1997. *Critique of Practical Reason*. Translated by Mary Gregor. edited by Mary Gregor. Cambridge: Cambridge University Press.

---. 1998. *Kant: Groundwork of the Metaphysics of Morals.Cambridge Texts in the History of Philosophy*, edited by Mary Gregor. Cambridge: Cambridge University Press.

Keller, Simon. 2004. "Friendship and Belief." *Philosophical Papers* 33 (3): 329-351.

Kelly, Thomas. 2002. "The Rationality of Belief and Other Propositional Attitudes." *Philosophical Studies* 110 (2): 163-96.

Kiesewetter, Benjamin. 2011. "'Ought' and the Perspective of the Agent." *Journal of Ethics and Social Philosophy* 5 (3): 1-24.

---. 2017. *The Normativity of Rationality*. Oxford University Press.

---. 2018. "What Kind of Perspectivism?" *Journal of Moral Philosophy* 15 (4): 415-443.

Kitcher, Philip. 1993. "Function and Design." *Midwest Studies in Philosophy* 18 (1): 379-397.

---. 2011. *The Ethical Project*. Cambridge, MA: Harvard University Press.

Klayman, Joshua, and Young-won Ha. 1987. "Confirmation, Disconfirmation, and Information in Hypothesis Testing." *Psychological Review* 94 (2): 211-228.

Korsgaard, Christine M. 1983. "Two Distinctions in Goodness." *Philosophical Review* 92 (2): 169-195.

---. 1996. *The Sources of Normativity*. Cambridge ; New York: Cambridge University Press.

---. 2009. "The Activity of Reason." *Proceedings and Addresses of the American Philosophical Association* 83 (2): 23 - 43.

Kusch, Martin. 2002. *Knowledge by Agreement : The Programme of Communitarian Epistemology: The Programme of Communitarian Epistemology*. Oxford: Oxford University Press.

---. 2009. "Testimony and the Value of Knowledge." In *Epistemic Value*, edited by Pritchard, Haddock and Millar, 60--94. Oxford: Oxford University Press.

---. 2011. "Knowledge and Certainties in the Epistemic State of Nature." *Episteme* 8 (1): 6-23.

Kusch, Martin, and Robin McKenna. 2020. "The Genealogical Method in Epistemology." *Synthese* 197 (3): 1057-1076.

Kvanvig, Jonathan L. 2003. *The Value of Knowledge and the Pursuit of Understanding*. Cambridge: Cambridge University Press.

---. 2008. "Pointless truth." *Midwest Studies in Philosophy* 32 (1): 199-212.

Langton, Rae. 2007. "Objective and Unconditioned Value." *Philosophical Review* 116 (2): 157-185.

Lewis, David K. 1970. "How to Define Theoretical Terms." *Journal of Philosophy* 67 (13): 427-446.

---. 1972. "Psychophysical and Theoretical Identifications." *Australasian Journal of Philosophy* 50 (3): 249-258.

---. 1973. *Counterfactuals*. Oxford: Blackwell.

Littlejohn, Clayton. 2019. "Being More Realistic About Reasons: On Rationality and Reasons Perspectivism." *Philosophy and Phenomenological Research* 99 (3): 605-627.

Loar, Brian. 1981. *Mind and Meaning*. Cambridge: Cambridge University Press.

Locke, John. 1988. *Two Treatises of Government*. edited by Peter Laslett. Cambridge: Cambridge University Press.

Lord, Errol. 2018. *The Importance of Being Rational*. Oxford: Oxford University Press.

Lycan, William G. 1996. *Consciousness and Experience*. Cambridge, MA: MIT Press.

Lynch, Michael P. 2004a. "Minimalism and the Value of Truth." *Philosophical Quarterly* 54 (217): 497 - 517.

---. 2004b. *True to Life: Why Truth Matters*. Cambridge, MA: MIT Press.

---. 2009. "The Value of Truth and the Truth of Values." In *Epistemic Value*, edited by A. Haddock, A. Millar and D. Pritchard, 225-242. Oxford: Oxford University Press.

Maguire, Barry. 2016. "The Value-Based Theory of Reasons." *Ergo: An Open Access Journal of Philosophy* 3.

Maguire, Barry, and Errol Lord. 2016. "An Opinionated Guide to the Weight of Reasons." In *Weighing Reasons*, 4-25.

Maguire, Barry, and Justin Snedegar. 2021. "Normative metaphysics for accountants." *Philosophical Studies* 178 (2): 363-384.

Maguire, Barry, and Jack Woods. 2020. "The Game of Belief." *Philosophical Review* 129 (2): 211-249.

Mandelbaum, Eric. 2014. "Thinking is Believing." *Inquiry: An Interdisciplinary Journal of Philosophy* 57 (1): 55-96.

Mantel, Susanne. 2017. "Three Cheers for Dispositions: A Dispositional Approach to Acting for a Normative Reason." *Erkenntnis* 82 (3): 561-582.

Marušić, Berislav. 2015. *Evidence and Agency Norms of Belief for Promising and Resolving*. Oxford: Oxford University Press.

Masip, Jaume, Eugenio Garrido, and Carmen Herrero. 2006. "Observers' Decision Moment in Deception Detection Experiments: Its Impact on Judgment, Accuracy, and Confidence." *International Journal of Psychology* 41 (4): 304-319.

Mason, H. E. 1996. *Moral Dilemmas and Moral Theory*. Oxford: Oxford University Press.

McGrath, Matthew. 2005. "Lynch on the Value of Truth." *Philosophical Books* 46 (4): 302-310.

McHugh, Conor. 2012. "The Truth Norm of Belief." *Pacific Philosophical Quarterly* 93 (1): 8-30.

---. 2013. "Normativism and Doxastic Deliberation." *Analytic Philosophy* 54 (4): 447-465.

McHugh, Conor, and Jonathan Way. 2017. "Objectivism and Perspectivism about the Epistemic Ought." *Ergo: An Open Access Journal of Philosophy* 4.

McHugh, Conor, and Daniel Whiting. 2014. "The Normativity of Belief." *Analysis* 74 (4): 698-713.

McKeever, Sean, and Michael Ridge. 2008. "Preempting principles: Recent debates in moral particularism." *Philosophy Compass* 3 (6): 1177-1192.

McNamara, Paul. 2010. "Deontic Logic." In *Stanford Encyclopedia of Philosophy*, edited by Ed Zalta.

Mele, Alfred R. 2001. *Self-Deception Unmasked*. Princeton, NJ: Princeton University Press.

Millikan, Ruth Garrett. 1984. *Language, Thought, and Other Biological Categories: New Foundations for Realism*. Cambridge, MA: MIT Press.

Mizrahi, Moti. 2009. "'Ought' Does Not Imply 'Can'." *Philosophical Frontiers* 4 (1): 19-35.

---. 2012. "Does 'Ought' Imply 'Can' from an Epistemic Point of View?" *Philosophia* 40 (4): 829-840.

Moran, Richard A. 2001. *Authority and Estrangement: An Essay on Self-Knowledge*. Princeton, NJ: Princeton University Press.

Nagel, Thomas. 1970. *The Possibility of Altruism*. Oxford: Clarendon Press.

---. 1989. "Fredom and the View from Nowhere." In *The View From Nowhere*, edited by Thomas Nagel. Oxford: Oxford University Press.

Neander, Karen. 1991. "The Teleological Notion of 'Function'." *Australasian Journal of Philosophy* 69 (4): 454 – 468.

Nichols, Shaun, and Stephen P. Stich. 2003. *Mindreading: An Integrated Account of Pretence, Self-Awareness, and Understanding Other Minds*. Oxford: Oxford University Press.

Nolfi, Kate. 2015. "How to be a Normativist about the Nature of Belief." *Pacific Philosophical Quarterly* 96 (2): 181-204.

Nozick, Robert. 1974. *Anarchy, State, and Utopia*. New York, NY: Basic Books.

O'Shaughnessy, Brian. 1963. "Observation and the Will." *Journal of Philosophy* 60 (14): 367-392.

Orwell, George 2003. *Nineteen Eighty-Four*. London: Penguin Books.

Owens, David J. 2003. "Does Belief Have an Aim?" *Philosophical Studies* 115 (3): 283-305.

Papineau, David. 2013. "There Are No Norms of Belief." In *The Aim of Belief*, edited by T. Chan, 64-79. Oxford: Oxford University Press.

Parfit, Derek. 2011. *On What Matters, 1.The Berkeley Tanner Lectures*. Oxford: Oxford University Press.

Peacocke, Christopher. 2008. *Truly Understood*. Oxford: Oxford University Press.

Pickard, Hanna. 2004. "Knowledge of Action Without Observation." *Proceedings of the Aristotelian Society* 104 (3): 203-228.

Piller, Christian. 2006. "Content-Related and Attitude-Related Reasons for Preferences." *Royal Institute of Philosophy Supplement* 59: 155-182.

Pittard, John, and Alex Worsnip. 2017. "Metanormative Contextualism and Normative Uncertainty." *Mind* 126 (501): 155-193.

Pollock, John, and Joe Cruz. 1999. *Contemporary Theories of Knowledge, 2nd Edition*. Lanham, ML: Rowman & Littlefield.

Pollock, John L. 1987. "Epistemic Norms." *Synthese* 71 (1): 61 - 95.

Porot, Nicolas, and Eric Mandelbaum. 2021. "The Science of Belief: A Progress Report." *WIREs Cognitive Science* 12 (2): e1539.

Pritchard, Duncan. 2016. "Epistemic Risk." *Journal of Philosophy* 113 (11): 550-571.

---. 2020. "Anti-risk Epistemology and Negative Epistemic Dependence." *Synthese* 197 (7): 2879-2894.

Pritchard, Duncan, Alan Millar, and Adrian Haddock. 2008. *Social Epistemology*. Oxford: Oxford University Press.

Proust, Marcel. 1996. *In Search of Lost Time Swann's Way*. London: Vintage

Queloz, Matthieu. 2019. "Genealogy and Knowledge-First Epistemology: A Mismatch?" *Philosophical Quarterly* 69 (274): 100-120.

---. 2020a. "From Paradigm-Based Explanation to Pragmatic Genealogy." *Mind* 129 (515): 683-714.

---. 2020b. "Revealing Social Functions through Pragmatic Genealogies." In *Social Functions in Philosophy: Metaphysical, Normative, and Methodological Perspectives*, edited by Rebekka Hufendiek, Daniel James and Raphael Van Riel, 200-218. London: Routledge.

Quilty-Dunn, Jake, and Eric Mandelbaum. 2018a. "Against Dispositionalism: Belief in Cognitive Science." *Philosophical Studies* 175 (9): 2353-2372.

---. 2018b. "Inferential Transitions." *Australasian Journal of Philosophy* 96 (3): 532-547.

Quine, Willard V. O. 1951. "Two Dogmas of Empiricism." *Philosophical Review* 60 (1): 20–43.

Rawls, John. 1955. "Two Concepts of Rules." *Philosophical Review* 64 (1): 3-32.

---. 1971. *A Theory of Justice*. Cambridge, MA: Harvard University Press.

Raz, Joseph. 2000. "The Truth in Particularism." In *Moral Particularism*, edited by Brad Hooker and Margaret Olivia Little, 48-78. Oxford: Oxford University Press.

---. 2010. "Reason, Reasons and Normativity." In *Oxford Studies in Metaethics, Volume 5*, edited by Russ Shafer-Landau. Oxford University Press.

Reisner, Andrew. 2009. "The Possibility of Pragmatic Reasons for Belief and the Wrong Kind of Reasons Problem." *Philosophical Studies* 145 (2): 257 - 272.

---. 2018. "Pragmatic Reasons for Belief." In *The Oxford Handbook of Reasons and Normativity*, edited by Daniel Star. Oxford University Press.

Rinard, Susanna. 2015. "Against the New Evidentialists." *Philosophical Issues* 25 (1): 208-223.

---. 2017. "No Exception for Belief." *Philosophy and Phenomenological Research* 94 (1): 121-143.

---. 2018. "Believing for Practical Reasons." *Noûs* (4): 763-784.

---. 2019. "Equal Treatment for Belief." *Philosophical Studies* 176 (7): 1923-1950.

Rosenthal, David M. 1986. "Two Concepts of Consciousness." *Philosophical Studies* 49 (May): 329-59.

Rousseau, Jean-Jacques. 1992. *Discourse on the Origin of Inequality*. Oxford: Oxford University Press.

Ryle, Gilbert. 1949. *The Concept of Mind*. Vol. 234. Chicago: University of Chicago Press.

Scanlon, T. M. 2014. *Being Realistic About Reasons*. Oxford: Oxford University Press.

Scanlon, Thomas. 1998. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press.

Schroeder, Mark. 2007. *Slaves of the Passions*. Oxford: Oxford University Press.

Schwenkler, John. 2011. "Perception and Practical Knowledge." *Philosophical Explorations* 14 (2): 137-152.

Schwitzgebel, Eric. 2001. "In-between Believing." *Philosophical Quarterly* 51 (202): 76-82.

---. 2002. "A Phenomenal, Dispositional Account of Belief." *Noûs* 36 (2): 249-275.

---. 2006. "Belief." In *Stanford Encyclopedia of Philosophy*.

---. 2010. "Acting contrary to our professed beliefs or the gulf between occurrent judgment and dispositional belief." *Pacific Philosophical Quarterly* 91 (4): 531-553.

---. 2013. "A Dispositional Approach to the Attitudes." In *New Essays on Belief*, edited by N. Nottelmann, 75-99. London: Palgrave.

Sellars, Wilfrid. 1969. "Language as Thought and as Communication." *Philosophy and Phenomenological Research* 29 (4): 506-527.

Sellars, Wilfrid S. 1956. "Empiricism and the Philosophy of Mind." *Minnesota Studies in the Philosophy of Science* 1: 253-329.

Setiya, Kieran. 2013. "Epistemic Agency: Some Doubts." *Philosophical Issues* 23 (1): 179-198.

Shah, Nishi. 2003. "How Truth Governs Belief." *Philosophical Review* 112 (4): 447-482.

---. 2006. "A New Argument for Evidentialism." *Philosophical Quarterly* 56 (225): 481–498.

Shah, Nishi, and J. David Velleman. 2005. "Doxastic Deliberation." *Philosophical Review* 114 (4): 497-534.

Shoemaker, Sydney. 1988. "On Knowing One's Own Mind." *Philosophical Perspectives* 2: 183-209.

---. 1994. "Self-knowledge and “Inner Sense": Lecture I: The Object Perception Model." *Philosophy and Phenomenological Research* 54 (2): 249-269.

Sider, Theodore. 2001. *Four Dimensionalism: An Ontology of Persistence and Time*. Oxford: Oxford University Press.

Siegel, Susanna. 2019. "Inference Without Reckoning." In *Reasoning: New Essays on Theoretical and Practical Thinking*, edited by Brendan Balcerak Jackson and Magdalena Balcerak Jackson, 15-31. Oxford: Oxford University Press.

Sinnott-Armstrong, Walter. 1984. "`Ought' Conversationally Implies `Can'." *Philosophical Review* 93 (2): 249-261.

Skorupski, John. 2007. "What is Normativity?" *Disputatio* 2 (23): 1 - 23.

---. 2010. *The Domain of Reasons*. Oxford: Oxford University Press.

Skurnik, Ian, Carolyn Yoon, Denise C. Park, and Norbert Schwarz. 2005. "How Warnings about False Claims Become Recommendations." *Journal of Consumer Research* 31 (4): 713-724.

Smith, Michael. 1994. *The Moral Problem*. Vol. 1. London: Blackwell.

Smithies, Declan. 2015. "Ideal Rationality and Logical Omniscience." *Synthese* 192 (9): 2769-2793.

Smyth, Nicholas. 2017. "The Function of Morality." *Philosophical Studies* 174 (5): 1127-1144.

Smyth, Nick. 2020. "A Genealogy of Emancipatory Values." *Inquiry: An Interdisciplinary Journal of Philosophy*: 1-39.

Snedegar, Justin. 2018. "Reasons for and Reasons against." *Philosophical Studies* 175 (3): 725-743.

Sorensen, Roy A. 1988. *Blindspots*. Oxford: Oxford University Press.

Sosa, Ernest. 2003. "The Place of Truth in Epistemology." In *Intellectual Virtue: Perspectives From Ethics and Epistemology*, edited by Linda Zagzebski and Michael DePaul, 155-180. Oxford: Oxford University Press.

Soteriou, Matthew. 2013. *The Mind's Construction: The Ontology of Mind and Mental Action*. Oxford: Oxford University Press.

Srinivasan, Amia. 2015. "Normativity Without Cartesian Privilege." *Philosophical Issues* 25 (1): 273-299.

Steglich-Petersen, Asbjørn. 2010. "The Truth Norm and Guidance: a Reply to Glüer and Wikforss: Discussions." *Mind* 119 (475): 749-755.

---. 2013. "The No Guidance Argument." *Theoria* 79 (1): 279-283.

Steglich-Petersen, Asbjørn, and Mattias Skipper.2019. "An Instrumentalist Account of How to Weigh Epistemic and Practical Reasons for Belief." *Mind* 129 (516): 1071-1094.

Strawson, Peter. 1962. "Freedom and Resentment." In *Proceedings of the British Academy, Volume 48: 1962*, 1-25.

Street, Chris N. H., and Daniel C. Richardson. 2015. "Descartes Versus Spinoza: Truth, Uncertainty, and Bias." *Social Cognition* 33 (3): 227-239.

Stroud, Sarah. 2006. "Epistemic Partiality in Friendship." *Ethics* 116 (3): 498-524.

Sylvan, Kurt. 2016a. "Epistemic Reasons I: Normativity." *Philosophy Compass* 11 (7): 364-376.

---. 2016b. "Epistemic Reasons II: Basing." *Philosophy Compass* 11 (7): 377-389.

Talbott, W. J. 1995. "Intentional Self-deception in a Single Coherent Self." *Philosophy and Phenomenological Research* 55 (1): 27-74.

Thomson, Judith Jarvis. 2008. *Normativity*. Chicago, IN: Open Court.

Tversky, Amos, and Daniel Kahneman. 1974. "Judgment under Uncertainty: Heuristics and Biases." *Science* 185 (4157): 1124-1131.

Unkelbach, Christian. 2007. "Reversing the Truth Effect: Learning the Interpretation of Processing Fluency in Judgments of Truth." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 33 (1): 219-230.

Van Bavel, Jay J., and Andrea Pereira. 2018. "The Partisan Brain: An Identity-Based Model of Political Belief." *Trends in Cognitive Sciences* 22 (3): 213-224.

Velleman, David. 2000a. "On the Aim of Belief." In *The Possibility of Practical Reason*, edited by David Velleman, 244-81. Oxford: Oxford University Press.

---. 2000b. *The Possibility of Practical Reason*. Oxford: Oxford University Press.

Von Hippel, William, and Robert Trivers. 2011. "The Evolution and Psychology of Self-deception." *Behavioral and Brain Sciences* 34 (1): 1.

von Wright, Georg Henrik. 1963. *Norm and Action: A Logical Enquiry*. London: Routledge and Kegan Paul.

Vranas, Peter B. M. 2007. "I Ought, Therefore I Can." *Philosophical Studies* 136 (2): 167-216.

Way, Jonathan. 2012. "Transmission and the Wrong Kind of Reason." *Ethics* 122 (3): 489-515.

---. 2016. "Two Arguments for Evidentialism." *Philosophical Quarterly* 66 (265): 805-818.

Way, Jonathan, and Daniel Whiting. 2016. "Reasons and Guidance." *Analytic Philosophy* 57 (3): 214-235.

---. 2017. "Perspectivism and the Argument from Guidance." *Ethical Theory and Moral Practice* 20 (2): 361-374.

Weatherson, Brian. 2014. "Running Risks Morally." *Philosophical Studies* 167 (1): 141-163.

Wedgwood, Ralph. 2002. "The Aim of Belief." *Philosophical Perspectives* 36 (s16): 267-97.

---. 2007. *The Nature of Normativity*. Oxford: Oxford University Press.

---. 2013a. "Doxastic Correctness." *Aristotelian Society Supplementary Volume* 87 (1): 217-234.

---. 2013b. "Rational 'Ought' Implies 'Can'." *Philosophical Issues* 23 (1): 70-92.

---. 2013c. "The Right Thing to Believe." In *The Aim of Belief*, edited by Timothy Chan, 123-139. Oxford University Press.

---. 2017. *The Value of Rationality*. Oxford: Oxford University Press.

Wei, Xintong. 2019. "The Permissible Norm of Truth and "Ought Implies Can"." *Logos and Episteme* 10 (4): 433-440.

---. 2020. "The Role of Pretense in the Process of Self-deception." *Philosophical Explorations* 23 (1): 1-14.

Whiting, Daniel. 2010. "Should I Believe the Truth?" *Dialectica* 64 (2): 213-224.

---. 2012. "Does Belief (Only) Aim at the Truth?" *Pacific Philosophical Quarterly* 93 (2): 279-300.

---. 2013. "Truth: the Aim and Norm of Belief." *Teorema: International Journal of Philosophy* 32 (3): 121-136.

---. 2017. "Against Second - Order Reasons." *Noûs* 51 (2): 398-420.

---. 2020. "Epistemic Worth." *Ergo: An Open Access Journal of Philosophy* 7.

---. 2021. "Aesthetic Reasons and the Demands They (Do Not) Make." *Philosophical Quarterly* 71 (2): 407-427.

Widerker, David. 1991. "Frankfurt on 'Ought Implies Can' and Alternative Possibilities." *Analysis* 51 (4): 222 - 224.

Williams, Bernard. 2002. *Truth and Truthfulness: An Essay in Genealogy*. Princeton, NJ: Princeton University Press.

Williams, Bernard 1973. *Problems of the Self*. Cambridge: Cambridge University Press.

---. 1981. *Moral Luck: Philosophical Papers, 1973-1980*. Cambridge: Cambridge University Press.

Williamson, Timothy. 2000. *Knowledge and its Limits*. Oxford: Oxford University Press.

---. 2007. *The Philosophy of Philosophy*. London: Wiley-Blackwell.

---. 2011. "Philosophical Expertise and the Burden of Proof." *Metaphilosophy* 42 (3): 215-229.

Wittgenstein, Ludwig. 1969. *On Certainty*. Vol. 323. Oxford: Blackwell.

Wright, Crispin. 1989. "Wittgenstein 's Later Philosophy of Mind: Sensation, Privacy, and Intention." *Journal of Philosophy* 86 (11): 622-634.

---. 1992. *Truth and Objectivity*. Cambridge, MA: Harvard University Press.

---. 2004. "Warrant for Nothing (and Foundations for Free)?" 78 (1): 167–212.

---. 2007. "Rule - following Without Reasons: Wittgenstein's Quietism and the Constitutive Question." *Ratio* 20 (4): 481–502.

Zagzebski, Linda. 2003. "The Search for the Source of Epistemic Good." *Metaphilosophy* 34 (1-2): 12-28.

Zahavi, Dan. 2005. *Subjectivity and Selfhood: Investigating the First-Person Perspective*. Vol. 3. Cambridge, MA: MIT Press.