OUGHTS, THOUGHTS, AND COMPANIONS IN GUILT

A DEFENSE OF MORAL REALISM

Zachary Mitchell Swindlehurst

39,491 Words

A thesis submitted in fulfillment of the requirements for the degree of

Master of Arts in Philosophy

The University of Otago

Dunedin, New Zealand

February, 2019

ABSTRACT

According to the *moral error theory*, there are no moral facts: all (positive, atomic) moral judgements are systematically and uniformly false. A popular strategy in recent years for arguing against the moral error theory is to deploy a *companions in guilt* (CG) argument. According to CG theorists, arguments for the moral error theory are insufficient, because either they rely on premises which do not warrant scepticism about moral facts, or they threaten to support an implausible error theoretic conclusion in other areas of discourse – areas which seem safe from such a conclusion. This thesis deploys a CG argument in order to defend *moral realism* – roughly, the view that moral judgements purport to state facts and that some of those judgements have true contents – against one influential argument for the error theory: J. L. Mackie's *argument from queerness*. The CG argument deployed depends on the assumption that *doxastic normativism* is true. Doxastic normativism is a metaphysical thesis according to which norms are in some sense constitutive of, or essential to, belief. Since my CG argument only works (if it works) on the assumption that doxastic normativism is true, much of this thesis is spent defending normativism against some of its influential detractors.

ACKNOWLEDGEMENTS

οὐκοῦν κἂν εἰ πρὸς αὐτὸ τὸ φῶς ἀναγκάζοι αὐτὸν βλέπειν, ἀλγεῖν τε ἂν τὰ ὄμματα καὶ φεύγειν ἀποστρεφόμενον πρὸς ἐκεῖνα ἃ δύναται καθορᾶν, καὶ νομίζειν ταῦτα τῷ ὄντι σαφέστερα τῶν δεικνυμένων;

And if someone compelled him to look at the light itself, wouldn't his eyes hurt, and wouldn't he turn around and flee towards the things he's able to see, believing that they're really clearer than the ones he's being shown?

– Plato, *Republic*, 515e

TABLE OF CONTENTS

INTRODUCTION

According to the *moral error theory*, there are no moral facts: positive, atomic moral judgements are systematically and uniformly false. According to *moral realism*, however, there are moral facts: moral judgements (e.g. "torturing children is wrong") purport to state facts and some such judgements have true contents. That is, moral judgements express beliefs which are apt for assessment in terms of truth and falsity, and some moral beliefs are indeed true. This thesis defends moral realism against one of the most influential arguments for the moral error theory – J. L. Mackie's *argument from queerness*. I do not attempt to provide a positive account of moral realism; I only hope to *defend* it against the error theorist.

In recent years, a popular strategy for arguing against the moral error theory is to deploy a *companions in guilt* (CG) argument. According to the CG theorist, arguments for metaethical anti-realism – such as the argument from queerness for the error theory, though some philosophers have developed CG arguments against other forms of anti-realism – are insufficient. Either they rely on premises which do not warrant scepticism about moral facts, or they threaten to support an implausible error theoretic conclusion in other areas of discourse – areas which appear safe from such a conclusion.

Likewise, in recent years, the philosophy of normativity has seen much interest in the metaphysical nature of belief. *Doxastic normativism* is the thesis that norms are (in a sense to be clarified) constitutive of, or essential to, belief, such that nothing not subject to those norms counts as a belief. This thesis attempts to develop a CG argument against the error theory based on doxastic normativism. Assuming normativism, I argue that if the argument from queerness were sound, then we could run the argument from queerness in the case of belief. A doxastic analogue of the argument from queerness would support the implausible conclusion that doxastic judgements (e.g. "Marco believes that Mallory reached the summit") are systematically and uniformly false. The upshot of this

conclusion would be the paradoxical one that there are *no beliefs and no believers*. This, I suggest, might constitute a *reductio ad absurdum* against the argument from queerness, because to commit to our doxastic analogue of the argument from queerness would be to commit *cognitive suicide.*

This thesis is structured as follows. Chapter 1 presents the threat to the error theorist: I argue that, assuming normativism, a doxastic analogue of the argument from queerness supports the conclusion that doxastic judgements are systematically and uniformly false. In §1.2, I contrast the error theory with moral realism, along with ethical non-cognitivism – another form of metaethical anti-realism. §1.3 outlines the argument from queerness as Mackie presents it. In §1.4, I give an overview of CG arguments. §1.4.1 presents a number of considerations in favor of what we can call *the parity premise* – the premise in CG arguments according to which moral facts and their companions stand or fall together. That is, the parity premise says that moral facts and doxastic facts (say) are sufficiently analogous for CG arguments to work: scepticism about the latter undermines scepticism about the former in such a way that warrants rejection of, for example, the moral error theory. §1.4.2 connects CG arguments with doxastic normativism, and clarifies what the latter amounts to. Although I do not defend normativism in this section, I argue that *if* normativism is true, then the argument from queerness would imply that nothing satisfies the norms which purport to be constitutive of belief. Finally, in §1.5, I flesh out a doxastic analogue of the argument from queerness. The conclusion of this argument – that doxastic judgements are systematically and uniformly false – has the paradoxical upshot there are no beliefs and no believers. This section concludes by considering (and rejecting) just one way in which a defender of the error theory might rebut our CG argument.

As will become clear at the end of Chapter 1, whether our CG argument works depends in large part on whether normativism is true. In Chapters 2 and 3, therefore, I defend normativism against some of its most influential detractors.

Chapter 2 considers the objections to normativism raised by Krister Bykvist and Anandi Hattiangadi. First, §2.2 further clarifies what doxastic normativism amounts to, following Paul Boghossian's presentation of an argument for it. In §2.3, I outline Bykvist and Hattiangadi's arguments against truth norms for belief. According to them, belief is not constitutively normative (at least when normativism is construed along the lines of a truth norm) because truth norms for belief cannot be formulated without unpalatable consequences. This we might call *the formulation problem*. In §2.4, I suggest we might rebut the formulation problem by holding that *knowledge*, rather than merely truth, is the aim and norm of belief. Adoption of a particular knowledge norm, I argue, allows us to preserve the virtues of the truth norm while avoiding its unpalatable consequences. Specifically, the knowledge norm does not succumb to the objections Bykvist and Hattiangadi raise against truth norms. Moreover, it is a plausible candidate for a constitutive doxastic norm – one that serves to distinguish belief from other cognitive attitudes. And finally, I argue that the knowledge norm can plausibly explain why the less controversial epistemic norms hold, since unlike the truth norm, the knowledge norm does not clash with the epistemic norms in some of the 'hard cases' (coin-toss and lottery cases, in particular).

Chapter 3 defends normativism against objections from Kathrin Glüer and Åsa Wikforss. According to their 'no guidance argument', if truth norms are constitutive of belief, then they must be capable of guiding belief formation; but if truth norms are in force, then they *cannot* guide belief formation, and so belief cannot be constitutively normative. Moreover, according to the 'regress of motivations' argument, not all belief can be rule- or norm-guided, otherwise we get a vicious infinite regress. (Although I defend the knowledge norm in Chapter 2, the no guidance argument is quite independent of the particular norm we endorse; so I assume the truth norm holds for most of Chapter 3.) In §3.2, I outline Glüer and Wikforss' no guidance argument and regress of motivations argument. §3.3 offers further clarifications of the no guidance argument, since it is

important to be clear about what the argument amounts to and a number of commentators have been misled by it. §3.4 suggests a potential way of blocking some of Glüer and Wikforss' arguments against normativism by an appeal to (a particular understanding of) the notion of *blind rule-following*. A foray into the literature on blind rule-following will be necessary in order to see how we might block some arguments against normativism. §3.4.1, therefore, outlines Boghossian's 'inference problem' for rule-following and Alex Miller's solution to it. In §3.4.2, I argue that Miller's solution to the inference problem might provide us with the tools to block some of Glüer and Wikforss' arguments against normativism. §3.5 considers a second way of blocking these arguments against normativism, from Teemu Toppinen. I argue that, although we might accept Toppinen's objections to Glüer and Wikforss, his own positive account of the way in which norms guide belief is likely to face difficulties. In §3.6, I conclude with a brief sketch of the way in which the knowledge norm defended in Chapter 2 might guide belief formation, though this account is necessarily incomplete and subject to a number of provisos.

1 AGAINST THE ERROR THEORY

1.1 INTRODUCTION

According to the moral error theorist, there are no moral facts: moral judgements are systematically and uniformly *false*. A popular strategy in recent years in arguing against the moral error theory is to develop a *companions in guilt* (CG) argument. According to CG arguments, arguments for the error theory are insufficient: either they rely on premises which do not warrant scepticism about moral facts, or they threaten to support an implausible error-theoretic conclusion in other areas of discourse – areas which seem safe from a sceptical conclusion, such as epistemic or practical reasoning. This chapter attempts to develop a CG argument against the error theory, connecting the literature on CG arguments (in the context of epistemic normativity) with the literature on *doxastic normativism*, the thesis that norms are in some sense constitutive of belief. In §1.2, I contrast the error theory with *moral realism*, and introduce what I call 'the problem of normativity' – the problem of explaining what moral facts could possibly be. §1.3 outlines *the argument from queerness*, J. L. Mackie's influential argument for the moral error theory. In §1.4, I give an overview of CG arguments, and clarifications of the sort of argument I will be making (§1.4.1). I argue that, *if* doxastic normativism is true and the argument from queerness goes through, then a doxastic analogue of the argument from queerness can be developed (§1.4.2). In §1.5, I flesh out this doxastic analogue of the argument from queerness. I argue that – assuming normativism – a doxastic analogue of the argument from queerness would be highly implausible. The conclusion of such an argument would be the paradoxical one that there are no beliefs and no believers; to believe it would be to commit *cognitive suicide*. The upshot is that if normativism is true, then the moral error theorist is in trouble, since a doxastic analogue of the argument from queerness (if sound) would prove too much: it would prove that there are no doxastic facts, i.e. no beliefs and no believers.

1.2 MORAL REALISM AND THE PROBLEM OF NORMATIVITY

Discussions of the metaphysics and epistemology of normativity abound in contemporary philosophy.[1] While ethics is perhaps the most obvious example of a normative area of thought and talk, normative concepts are numerous and varied. On the face of it, a concept is normative when either certain 'oughts' or obligations can be derived from judgements involving it, or when prior norms are in force that govern the application of the concept. Thus it has become standard in recent discourse for philosophers to proceed from the assumption that, *inter alia*, concepts such as *virtue*, *justice*, *meaning*, *mental content*, and *knowledge* are all intrinsically normative concepts.[2] Take virtue, for example. According to Rosalind Hursthouse, the concept of a virtue "is the concept of something that makes its possessor good; a virtuous person is a morally good, excellent, or admirable person who acts and reacts well, rightly, as she should" (2000: 13). Virtue, therefore, is a normative concept in that a virtuous agent is someone who has, in the course of becoming virtuous, incurred a certain disposition to act as she *ought*. So the key feature of normative concepts seems to be that they are in some sense *prescriptive* or *action-guiding*. This chapter draws a parallel between moral normativity and *doxastic* normativity. I shall say more about the sense in which belief purports to be normative below, as well as in Chapters 2 and 3.

Some philosophers have claimed that a significant ontological challenge is borne by the moral realist: the challenge of explaining how there can possibly *be* normative facts of the sort required for the theory to be true. Call this *the problem of normativity*. Moral realism, in the sense discussed here, is the following view (Sayre-McCord 1986):

---

[1] Cf. Finlay (2010).
[2] See Korsgaard (1996) for discussion.

(MORAL REALISM): Moral judgements (i) purport to state facts; and (i) some

moral judgements are true.[3]

For example, according to the moral realist, the moral judgement "torturing children is wrong" expresses the *belief*, apt for assessment in terms of truth and falsity, that torturing children is wrong. Moreover, the judgement has the truth-value *true* if and only if it is the case that torturing children is wrong. Moral realists are therefore *cognitivists* about moral judgement, i.e. the judgements in question express truth-apt beliefs. By contrast, *non-cognitivists* about moral judgement deny in the first place that moral judgements even purport to state facts; rather, the judgements in question express *conative* (rather than *cognitive*) mental states such as feelings, sentiments, or attitudes of approval and disapproval.[4] Moral judgements, according to the non-cognitivist, are not even truth-apt, so the moral judgement "torturing children is wrong" just expresses something like an attitude of disapproval towards torturing children.

Moral realism is opposed by both *cognitivist* and *non-cognitivist* anti-realism. Cognitivist anti-realists accept (i) but deny (ii); non-cognitivists deny both (i) and (ii). The moral realist appears to face an explanatory burden not shared by anti-

---

[3] Of course, moral realism as understood here is very broad. In particular, (MORAL REALISM) as defined here says nothing about whether moral facts need to be mind-independent. Thus even Korsgaard's (1996) Kantian constructivism, according to which moral facts are mind-dependent, would on the present picture count as moral realism. This might be controversial, but nothing in what follows will turn on this.

[4] The assumption underlying (our present understanding of) the distinction between cognitivism and non-cognitivism, then, is this: the cognitivist endorses a *factualist* moral semantics which goes hand-in-hand with their psychological account of moral judgement. That is, in saying that moral judgements are truth-apt, the cognitivist is saying that those judgements express beliefs. On the other hand, the non-cognitivist endorses a *non-factualist* moral semantics which goes hand in hand with their account of moral judgement. That is, in saying that moral judgements are not truth-apt, the non-cognitivist is saying that moral judgements express some conative mental state which lacks truth-conditions. However, it is worth noting that some philosophers separate their account of moral semantics from their account of moral psychology. For example, Richard Joyce (2001) combines *factualism* about moral semantics with *cognitivism* about moral psychology. While he maintains that cognitivism gives an accurate description of our current moral practice, he advocates a *revolutionary moral fictionalism* involving a *revisal* of our moral practice along non-cognitivist lines. By contrast, Mark Kalderon's *hermeneutic moral fictionalism* (2005) combines *factualism* about moral semantics with *non-cognitivism* about moral psychology. See §6.7 and §6.10 of Miller (2013) for a discussion of these issues.

realists of both sorts: what could it possibly *be* that makes moral claims true? The moral realist is committed to the existence of moral facts, and so must give an account of their metaphysics and epistemology.[5] J.L. Mackie (1977) is an ethical sceptic who retains a commitment to cognitivism; he argues for an error theory about moral judgement, according to which moral judgements do purport to state facts, only they are systematically and uniformly false.[6,7] The moral judgement "torturing children is wrong", then, does express the belief that torturing children is wrong, only it is false, since there is nothing in the world which acts as its truth-maker. On the other hand, early ethical non-cognitivists, such as A. J. Ayer (1946), argue for an emotivist theory of moral language, according to which moral judgements express sentiments or attitudes of approval and disapproval.[8] This is sometimes – though often pejoratively – labeled the "boo!/hurrah!" theory of moral judgement: to say that torturing children is wrong is like saying "Boo! (torturing children)"; to say that one ought to give to charity is like saying "Hurrah! (giving to charity)". This thesis focuses primarily on the error theory. A *full* defense of moral realism would need to rebut non-cognitivism, but I can only attempt to rebut Mackie's arguments here. If we can show that the

---

[5] Note that some moral realists – often those of a broadly Wittgensteinian bent – have a prior commitment to metaphysical or philosophical *quietism*, and so will feel no pressure to give such an account of the metaphysics and epistemology of moral facts. Non-reductionists such as John McDowell (1994, 1998), for instance, take normativity to be of a more primitive character and are unlikely to accept the explanatory burden being demanded here. I will not address such positions at present; rather, I want to *concede* that the explanation of normativity poses a (potential) problem for moral realism and assess the implications of this for other kinds of realism.

[6] More precisely, the error theorist is committed to the claim that all *positive, atomic* moral judgements (e.g., "One ought to give to charity") are systematically and uniformly false (cf. Pigden 2007). Nothing in what follows will turn on this, so I will not use the 'positive, atomic' qualification hereafter.

[7] Mackie is certainly not the *only* error theorist, nor is the argument from queerness the only argument for the error theory. I focus on Mackie, since the argument from queerness is by far the most influential argument for the moral error theory. However, see Hinckfuss (1987), Garner (1990, 2007), Joyce (2001), Olson (2014), and Streumer (2017). 'The error theory' hereafter refers primarily to Mackie's version of it, unless otherwise specified.

[8] See also Stevenson (1944) and Hare (1952) for classic defenses of ethical non-cognitivism. Contemporary defenders of ethical non-cognitivism (though they eschew the label "non-cognitivism" in favor of "expressivism") include Blackburn (1984) and Gibbard (1990).

argument from queerness is unsound, then we will have taken some first steps towards solving the problem of normativity.

1.3 THE ARGUMENT FROM QUEERNESS

Let us examine Mackie's argument for the moral error theory: the argument from queerness. First, however, we need to examine Mackie's prior, conceptual commitments. According to Michael Smith (1994: Chapter 3), what allows Mackie to argue for the error theory is a conceptual commitment to *moral rationalism*, or strong *motivational internalism*. Roughly put, moral rationalism is the thesis that there is an internal relation between what is *morally right* and agents' reasons for action; and motivational internalism, roughly put, is the thesis that there is an internal relation between *moral judgement* and motivation. More precisely, Smith distinguishes three versions of the thesis, the first of which entails either the second or the third:

(MORAL RATIONALISM): It is a conceptual truth that if it is right for agents to φ in circumstances C, then there is a reason for those agents to φ in C.

(MOTIVATIONAL INTERNALISM*): It is a conceptual truth that if an agent judges it right for her to φ in circumstances C, then she is motivated to φ in C.

(MOTIVATIONAL INTERNALISM**): It is a conceptual truth that if an agent judges it right for her to φ in circumstances C, then either she is motivated to φ in C or she is practically irrational.

Thus commitment to (MORAL RATIONALISM) is quite a strong commitment, while commitment to either form of motivational internalism on its own is a weaker

commitment.[9] (MORAL RATIONALISM) makes no mention of the agents' *judgements* that φ-ing is right: if vegetarianism is morally right, then an agent *A* has a reason not to eat meat. (MOTIVATIONAL INTERNALISM*) is weaker – it mentions *A*'s judgement that φ-ing is right – but it is still fairly strong since it carries motivation with it *simpliciter*: if *A* judges vegetarianism to be right, her motivation not to eat meat cannot be defeated by such commonplace phenomena as weakness of the will. So while (MOTIVATIONAL INTERNALISM*) is weaker than (MORAL RATIONALISM), it is still fairly strong. (MOTIVATIONAL INTERNALISM**) puts a *practicality requirement* on moral judgement. It is *defeasible* in that it can accommodate weakness of the will and other motivation-sapping phenomena: if *A* judges vegetarianism to be right, then *either* she is motivated not to eat meat *or* she is practically irrational. Someone is practically irrational when they are suffering from weakness of the will or other sorts of practical akrasia that sap one's motivation; if someone is depressed, for example, then we cannot expect their moral reasons to line up with the course of action (or lack of action) that they take. So, *prima facie*, (MOTIVATIONAL INTERNALISM**) seems to be the most plausible commitment out of the three.[10]

It is important to emphasize that, for the rationalist and motivational internalist, (MORAL RATIONALISM) and both versions of motivational internalism are *conceptual truths*. Motivational *externalists* will accept as a *contingent fact* about human beings that, say, if *A* judges vegetarianism to be right, then either she is motivated not to eat meat or she is practically irrational. The difference is the

---

[9] It is important to note that while (MORAL RATIONALISM) *entails* whichever version of motivational internalism is true, one can be committed to motivational internalism *without* being committed to moral rationalism. Since moral rationalism entails (some version of) motivational internalism, someone who thinks that motivational internalism is false will have to say that moral rationalism is false by *modus tollens*.

[10] One might worry, however, that (MOTIVATIONAL INTERNALISM**) is *too* weak. As Anandi Hattiangadi puts it: "if one can make a normative judgement and still not be motivated (because one is practically irrational), then it seems as though there is no longer a conceptual or necessary link between normative judgements and motivation – normative judgements do not introduce a 'new relation'" (2007: 42). I cannot explore these issues properly here. As such, I will assume that some version of motivational internalism is plausible.

externalist denies that this is a conceptual truth.[11] On their view, it is just a contingent fact about human beings that, barring any practical irrationality, they are motivated to act in accord with their moral judgements; there is no *conceptual* connection between moral judgement and motivation.

Now, let us simply grant Mackie his commitment to (MORAL RATIONALSIM).[12] Smith points out that in order for the argument from queerness to get off the ground, Mackie must ask both a *conceptual* question and a *substantive* or *ontological* question about the nature of putative moral facts. Consider an analogous case. Suppose we are interested in the question of whether there are any unicorns. In order to answer this question, we need to ask both a conceptual and a substantive/ontological question about the nature of unicorns. First, we ask what our *concept* of a unicorn is. We might say that our concept of a unicorn is the concept of a horse-like creature with a magical horn. Second, we ask whether there is anything in the world that instantiates that concept; that is, we ask whether there are really any horse-like creatures with magical horns. Since our answer to the second question will be *no*, we will say that there are no unicorns.

Mackie follows essentially the same procedure. In order to answer the question, "Are there any moral facts?" (or "objective values" as Mackie puts it), we need to first ask what our concept of a moral fact is, and second, whether there is anything in the world instantiating that concept. In answering the conceptual question, Mackie makes use of Kant's distinction in the *Groundwork of the Metaphysics of Morals* between *hypothetical* imperatives, on the one hand, and *categorical* imperatives on the other.[13] Hypothetical imperatives are imperatives that can be overridden by a change in one's contingent desires; they are, in

---

[11] See, e.g., Sturgeon (1985), Railton (1986), and Brink (1986); cf. Foot (1972).

[12] Many philosophers who reject the error theory reject it precisely because they think this commitment is too strong. See, e.g., Finlay (2008), and Joyce (2011) for a reply to such criticism.

[13] Kant (1785: §II). For example: "[A]ll *imperatives* command either *hypothetically* or *categorically*. The former represent the practical necessity [*praktische Notwendigkeit*] of a possible action as a means to achieving something else that one wants (or at least that is possible for one to want). The categorical imperative would be the one that represented an action as objectively necessary [*objectiv-notwendig*] by itself, without reference to another end" (§II, 4.414, 57).

Kant's terminology, "practically necessary" (*praktisch-notwendig*) and take the general form: "If you want *X*, then you ought to do *Y*". But categorical imperatives have normative force independently of our desires; they are "objectively necessary" (*objectiv-notwendig*) and take the general form: "You ought to do *Y*".[14] A categorical imperative to do *Y* would hold regardless of whether one desired *X*. For example, "Torturing children is wrong" contains a categorical imperative not to torture children; it holds regardless of any desire to (not) torture children. By contrast, the *ought* in the imperative, "If you want to get to the Octagon by 5:30pm, then you ought to take the 5:00pm bus" is merely a hypothetical *ought*; its normative force comes entirely from the contingent desire you have to get to the Octagon by 5:30pm – if you have no desire to get to the Octagon by 5:30pm, then you have no reason to take the 5:00pm bus.

Now, for Mackie, our concept of a moral fact is the concept of a fact about an *objectively prescriptive* feature of the world, where an objectively prescriptive feature of the world is a *categorical reason for action*.[15] Recall Mackie's commitment to moral rationalism. If it is morally right to *φ*, then agents have a reason to *φ*. "A reason" here should be read as "a *categorical* reason": for Mackie, categorical imperatives of the form "You ought to do *Y*" just say that there is a categorical reason to do *Y*, independently of any contingent desires for *X*. So Mackie has answered the conceptual question: our concept of a moral fact is the concept of an objectively prescriptive feature of the world, or a fact about a categorical

---

[14] The distinction is *not* merely grammatical; the point is not that hypothetical imperatives are grammatically conditional, while categorical imperatives are not. Rather, it is that categorical imperatives cannot be overridden by a change in one's contingent desires. Indeed, categorical imperatives can sometimes be conditional in form: "If you see a drowning child, you ought to save him" is presumably a categorical imperative (if there are any categorical imperatives), because no facts about your contingent desires bear on whether you ought to save the child.

[15] I will assume in what follows that the relationship between (normative) *reasons* and *oughts* is this: reasons are facts that *count in favour* of some performance and explain what we ought to do (cf. Raz 1975). This assumption is not entirely uncontroversial (see, e.g., Railton 1986; Broome 2000, 2004). However, it should be relatively harmless in the present context. Below, I draw a parallel between *moral* reasons or oughts and *doxastic* reasons or oughts. For this parallel to work, all we need is for belief to be categorically normative, regardless of whether the normativity in question is construed in terms of reasons or in terms of oughts. After all, Broome himself acknowledges that our most familiar normative concepts are *ought* and *a reason* (2000: 78).

moral reason for action (*pace* Foot 1972). But what of the substantive/ontological question? Mackie's answer is that moral facts, if they existed, would be metaphysically and epistemologically 'queer':

> If there were objective values, then they would be entities or qualities or relations of a very strange sort, utterly different from anything else in the universe. Correspondingly, if we were aware of them, it would have to be by some special faculty of moral perception or intuition, utterly different from our ordinary ways of knowing anything else (1977: 38).

On the metaphysical side, moral facts would have to be *sui generis*, akin to Moore's non-natural qualities (Moore 1903) or Plato's Form of the Good. On the epistemological side, knowledge of the 'authoritative prescriptivity' of moral facts would have to come from some special faculty of moral perception or intuition. If we had knowledge of some *distinctively ethical* premise, for example, the way we would have to have acquired this knowledge would be by some special sort of intuition: none of our ordinary ways of knowing (sensory perception, introspection, confirmation of some explanatory hypothesis, and so on) would do. So objectively prescriptive features of the world would be epistemologically queer too. Moreover, according to Mackie, arguments for the existence of moral facts such as those of Plato and Moore provide us with a "dramatic picture" of what such facts would have to be:

> The Form of the Good is such that knowledge of it provides the knower with both a direction and an overriding motive; something's being good both tells the person who knows this to pursue it and makes him pursue it. An objective good would be sought by anyone who was acquainted with it, not because of any contingent fact that this person, or every person, is so constituted that he desires this end, but just because the end has to-be-pursuedness somehow built into it. Similarly, if there were objective principles of right and wrong, any wrong (possible) course of action would have not-to-be-doneness somehow built into it (1977: 40).

Thus if there were any objectively prescriptive features of the world, they would be metaphysically and epistemologically queer: they would provide one with a categorical moral reason for action or abstention, but it is not clear what such

things could possibly be or how we could possibly know them. The upshot of this, according to Mackie, is that there are *no* objectively prescriptive features of the world – i.e., there are no categorical moral reasons for action – so there are no moral facts. We can therefore (following Miller 2010) reconstruct the argument from queerness as follows:

(1) Our concept of a moral fact is a concept of a reason for action.

(2) If our concept of a moral fact is a concept of a reason for action, then our concept of a moral fact is a concept of an objectively prescriptive feature of the world.

(3) Therefore, our concept of a moral fact is a concept of an objectively prescriptive feature of the world.

(4) But if our concept of a moral fact is a concept of an objectively prescriptive feature of the world and there are no objectively prescriptive features of the world, then moral judgements are systematically and uniformly false.

(5) There are no objectively prescriptive features of the world.

(6) Therefore, moral judgements are systematically and uniformly false.

Subconclusion (3) follows straightforwardly from premises (1) and (2). And we can grant premise (4): it is obvious that moral judgements would be systematically and uniformly false if we accept the conceptual claim that moral facts would have to be objectively prescriptive but deny the ontological claim that there are any objectively prescriptive features of the world. (By analogy, our unicorn judgements are systematically and uniformly false because there is nothing in the world that instantiates our concept of a unicorn.) So if – contra the moral externalists – we can grant premise (1), then the argument turns on premise (5): we can get to the main conclusion (6) only if there really are no objectively prescriptive features of the world. But we saw that, according to Mackie, there can

be no such features; for if there were, they would be metaphysically and episte-mologically queer. So there are no moral facts: moral judgements are systematically and uniformly false.

For our present purposes, let us suppose that the argument from queerness is sound or at least plausible – that is, suppose Mackie's argument poses a significant challenge to the moral realist (or at least to the moral realist who accepts premise (1)). The realist, then, faces the challenge of explaining what moral facts – if they exist – could possibly *be*.

1.4 COMPANIONS IN GUILT ARGUMENTS

A popular strategy in recent literature for arguing against the moral error theory (or sometimes against other forms of metaethical anti-realism, such as non-cognitivism) is to develop a 'companions in guilt' (CG) argument, according to which the argument from queerness is insufficient because it threatens to undermine the idea that there are categorical epistemic or practical norms.[16] The idea is that there are objectively prescriptive epistemic facts (or whatever), and since there is nothing 'queer' about *these* facts, there shouldn't be anything queer about moral facts either. In other words, CG arguments identify an area of discourse which involves categorical norms but which is 'safe' from the argument from queerness. Mackie himself suggests that this is the most promising strategy for rebutting the argument from queerness:

> [T]he best move for the moral objectivist is not to evade [the argument from queerness], but to look for companions in guilt. For example, Richard Price argues that it is not moral knowledge alone that such an empiricism as those of Locke and Hume is unable to account

---

[16] Perhaps one of the most influential CG arguments, based on the parallel between moral and epistemic normativity, is due to Cuneo (2007). See also Stratton-Lake (2002: xxv-xxvi), Kelly (2003), Shafer-Landau (2003: 113), Rowland (2013, 2016), Greco (2015), Das (2016, 2017), Worsnip (2016), and Clipsham (201X). Heathwood (2009), Olson (2011), Husi (2013) and Cowie (2014, 2016) offer criticism of this sort of argument. Note that CG arguments need not necessarily appeal to epistemic normativity: other CG arguments draw a parallel with mathematics (Clark-Doane 2012), prudence (Fletcher 2017), love (Keller 2017), or *any* type of normative judgment (Bedke 2010; cf. Streumer 2017). See also Lillehammer (2007) for a comprehensive, critical survey of the literature on CG arguments.

for, but also our knowledge and even our ideas of essence, number, identity, diversity, solidity, inertia, substance, the necessary existence and infinite extension of time and space, necessity and possibility in general, power, and causation. If the understanding, which Price defines as the faculty within us that discerns truth, is also a source of new simple ideas of so many other sorts, may it not also be a power of immediately perceiving right and wrong, which yet are real characters of actions? […] This is an important counter to the argument from queerness. The only adequate reply to it would be to show how, on empiricist foundations, we can construct an account of the ideas and beliefs and knowledge that we have of all these matters (1977: 39).

The idea behind "the only adequate reply" to this objection is that, for example, necessity and possibility might seem 'queer', and so in order to avoid an error theoretic conclusion about necessity and possibility, one would need to give an empiricist account of them. It is perhaps surprising that Mackie does not pursue this counter to the argument from queerness much further. Indeed, Mackie is clear that the argument from queerness does not undermine the objectivity of moral values alone:

> The claim that values are not objective, are not part of the fabric of the world, is meant to include not only moral goodness, which might be most naturally equated with moral value, but also other things that could be more loosely called moral values or disvalues – rightness and wrongness, duty, obligation, an action's being rotten and contemptible, and so on. It also includes non-moral values, notably aesthetic ones, beauty and various kinds of artistic merit (1977: 15).

But why stop at aesthetic value? This thesis discusses *doxastic* value and normativity: if there are objective or categorical doxastic norms, then the argument from queerness – if sound – will surely undermine them. Before we get into this, however, it is worth saying something more about precisely the sort of argument I shall be making.

16

1.4.1 HOW COMPANIONS IN GUILT ARGUMENTS WORK

In a comprehensive survey of the literature on CG arguments, Hallvard Lillehammer (2007) distinguishes two kinds of CG argument: those by *entailment*, on the one hand, and those by *analogy*, on the other. Lillehammer gives the following characterization of CG arguments by entailment:

> Companions in guilt arguments by entailment work by showing that some set of philosophically problematic claims are necessarily implied by a different set of claims that are agreed not to be philosophically problematic in the same way […] Arguments by entailment in ethics and the philosophy of value relate ethical or other evaluative claims to some different set of claims, the objectivity of which is not in serious doubt, and then purport to show that the relevant ethical or evaluative claims actually follow by necessity from the truth of these underlying claims. With respect to their truth and objectivity, the two sets of claims are companions in guilt as a matter of necessity (2007: 11).

Lillehammer notes that entailment is not a matter of degree: the idea is that scepticism about ethical value *necessarily* implies scepticism about its companion, in an 'all or nothing' sort of way. By contrast, CG arguments by analogy identify companions which might in some respects be similar to ethical values and dissimilar in other respects, where similarity *is* a matter of degree:

> Arguments by analogy compare two sets of claims with the aim of showing that they are similar in some important respect […] Companions in guilt arguments by analogy in ethics and the philosophy of value compare evaluative claims with some set of claims generally assumed to be less problematic with respect to their objectivity, and then show that the two sets of claims exhibit the very same features that, in the evaluative case, have generated scepticism about their objectivity. With respect to the features in question, the two sets of claims are companions in guilt (2007: 13).

In what follows, I focus primarily on CG arguments by analogy. Indeed, the argument I will be making relies on the claim that there is sufficient similarity

between moral and doxastic normativity in order for the argument from queerness to be run against the latter.[17] There are two issues here. First, we need to ask whether there is sufficient 'parity' between moral and doxastic facts for this strategy to work. Second, we need an account of the doxastic norms: are they *categorical*, and if so, what grounds them? These two issues correspond to two ways a defender of the error theory might block CG arguments. For one thing, the error theorist might deny that moral reasons and doxastic reasons are sufficiently analogous: she might, for instance, identify some key respect in which moral norms and doxastic norms are dissimilar, thereby blocking the claim that they stand or fall together. Moreover, the error theorist might deny that doxastic norms are categorical: she might say that the doxastic *ought* is a mere hypothetical *ought* that can be 'translated away' so as to give a naturalistic account of it.[18]

Let us start with the first issue. To begin with, take two influential examples of CG arguments by analogy, based on a parallel between moral and *epistemic* normativity. The first is Terence Cuneo's 'core argument' for moral realism (2007: 6):

(i) If moral facts do not exist, then epistemic facts do not exist.

---

[17] Many extant CG arguments rely on a parallel between moral and *epistemic* normativity. My own CG argument, however, relies on a parallel between moral and *doxastic* normativity. I discuss epistemic normativity in the context of CG arguments only for illustrative purposes. But what is the difference between doxastic and epistemic normativity? Epistemic norms are norms the satisfaction of which is necessary in order for a subject to have *justified* beliefs. By contrast, as will become clear below, doxastic norms are the norms necessary in order for a subject to have *beliefs* proper. Similarly, *epistemic reasons* are reasons for belief, whereas *doxastic reasons* are reasons for *ascription* of belief. The distinction is subtle, but important, since my own CG argument (based on doxastic normativity) will if sound provide us with a different way of rejecting the argument from queerness, from that of many extant CG arguments.

[18] Indeed, the hypothetical imperative "If you want to get to the Octagon by 5:30pm, then you ought to take the 5:00pm bus" can be translated away as: "Taking the 5:00pm bus will get you to the Octagon by 5:30pm". That is why hypothetical imperatives are generally thought to square neatly with naturalism. I cannot properly explore the issue of normativity vs naturalism here. For some classic discussions of the matter, see Hume on 'is' and 'ought' (1738: §3.1.1.27), Moore on the open-question argument (1903: Chapter 1), and Sellars on the distinction between the 'causal space of nature' and the 'normative space of reasons' (1956). See also Darwall, Gibbard, and Railton (1992) for useful discussion.

(ii) Epistemic facts exist.

(iii) Therefore, moral facts exist.

(iv) If moral facts exist, then moral realism is true.

(v) Therefore, moral realism is true.

The second influential CG argument by analogy is Richard Rowland's 'argument from epistemic reasons' (2013: 1):

(i) According to moral error theory, there are no categorical normative reasons.

(ii) If there are no categorical normative reasons, then there are no epistemic reasons for belief.

(iii) But there are epistemic reasons for belief.

(iv) Therefore, there are categorical normative reasons.

(v) Therefore, moral error theory is false.

It should be relatively clear – at least assuming there are epistemic facts/reasons for belief – that both arguments turn the same respective premise, (i) in Cuneo's core argument, and (ii) in Rowland's argument from epistemic reasons. Whether the CG argument goes through depends on the parity between moral facts and epistemic facts. The idea behind this is that epistemic reasons, like moral reasons, purport to be *categorically normative*: for example, you ought (categorically) to form beliefs in accordance with evidence. If epistemic reasons are indeed categorically normative and they exist, then the argument from queerness is unsound, since it relies on the premise the there are *no* categorically normative properties or objectively prescriptive features of the world (premise (5) in the argument from queerness above). Call the premise according to which moral and epistemic norms stand or fall together *the parity premise*. (I take it that our considerations of the parity premise regarding moral and epistemic normativity will apply *mutatis mutandis* to doxastic normativity, if we can show that doxastic norms are categorically normative.)

Is the parity premise true? Suppose – at least for the moment – that the error theorist concedes that epistemic facts are categorical. Given this supposition, one worry might be that although moral and epistemic reasons are alike with respect to their categoricity, they are unlike in another respect: belief is about *facts*, ethics is about *acts*. In other words, moral reasons are reasons to perform some act, whereas epistemic reasons are reasons for belief. Thus there is a disanalogy between moral and epistemic reasons, and this counts against the parity premise being true. But this objection seems to miss the mark. As Rowland argues:

> But moral error theorists are not, *qua* error theorists, skeptics about the facts (or propositions) that we take to be moral reasons. Nor are they skeptics about the acts for which we take these facts to be reasons. Rather, moral error theorists are skeptics about the reason relation that we take to hold between the facts and the acts. Some fact's being a normative reason for an act is just its having the relational property of being a normative reason for an act […] Moral error theorists claim that categorical normative warranting relations do not exist. But if the moral error theorists' skepticism is about categorical normative relations, they cannot be skeptical about categorical normative relations that have acts as one of their relata, but not about categorical normative relations that have beliefs as one of their relata (2013: 3-4).

The idea is that, while it is true that in the moral case the reasons in question are reasons for *action*, and that in the epistemic case the reasons in question are reasons for *belief*, this does not undermine the parity premise. For we can think of reasons for belief and reasons for action as different relata of the same normative relation: that of *being a reason for*.[19] (However, it seems correct to say that reasons' being a kind of warranting relation does not preclude them from being facts, for relations can also be facts.)

---

[19] According to Alan Millar (2009), the parity between moral and epistemic reasons is this: both reasons for intentional action and reasons for belief are defined by a *constitutive aim*, though they differ in what that aim is. Whereas belief is defined by the constitutive aim of truth, "the constitutive aim of intentional action is simply that the action should have a point, as being either a means to an end or the realization of some end of the agent" (2009: 143). I cannot properly pursue this line of argument here, though I discuss belief's having an aim below.

It has been suggested that there is a more fundamental issue with CG arguments. By way of criticism, Christopher Cowie reconstructs a *general* CG argument by analogy as follows (2014: 408):

(i) If the arguments for the moral error theory are sufficient to establish its truth, then those arguments (or appropriate analogues of them) are also sufficient to establish the truth of the epistemic error theory. (*Parity premise*.)

(ii) The epistemic error theory is false. (*Epistemic existence premise.*)

(iii) Therefore, the arguments for the moral error theory are not sufficient to establish that the moral error theory is true.

Cowie argues that there is an internal tension between the parity premise (i) and the epistemic existence premise (ii) of the general CG argument. A consequence of this, he claims, is that CG arguments *cannot* work: the internal tension between (i) and (ii) undermines either the *soundness* or the *dialectical force* of CG arguments. Cowie (2014: 410) gives two objections to the idea that premises (i) and (ii) of the general CG argument yield the conclusion (iii). On the one hand, there is *the objection from disparity*, according to which the arguments typically offered for the epistemic existence premise undermine the parity premise. On the other hand, there is *the objection from redundancy*, according to which there are *some* arguments for the epistemic existence premise which do not undermine the parity premise, but which nevertheless render CG arguments dialectically redundant. If either of these objections succeed, then they undermine either the soundness or the dialectical force of CG arguments, respectively. Let us examine each of Cowie's objections in turn.

To begin with, note that Cowie's rebuttal of CG arguments takes the form of a dilemma. The idea is that affirming the reality of epistemic reasons does not undermine scepticism about moral reasons, because:

(a) *Either* the objection from disparity undermines the parity premise and hence the companions in guilt argument;

(b) *Or* the objection from redundancy renders the companions in guilt argument dialectically redundant (Cowie 2014: 410).

After fleshing out each objection, Cowie draws an ambitious conclusion from this dilemma: if there is an internal tension between (i) and (ii) such that we get a dilemma of the sort Cowie introduces, then CG arguments *could not possibly work*. Let us examine, then, both horns of the dilemma.

Let us start with the first horn (a). According to Cowie's disparity objection, many arguments typically offered for the epistemic existence premise end up undermining the parity premise. But what are these arguments for the epistemic existence premise? According to one influential line of argument, the epistemic existence premise must be true because its denial would be *self-defeating*.[20] Call this *the argument from self-defeat*. Here is Philip Stratton-Lake:

> If [...] value is to be understood in terms of reasons, then (assuming one has no problem with actions and psychological attitudes) one's scepticism about value would have to stem from scepticism about reasons. But if an error theory about value rests on an error theory about reasons, then the error theorist would be committed both to the view that we have good reason to deny the existence of reasons, and that this implies that all of our reasons judgements are false. But such a view seems to undermine itself, for it is to say that we have reason to be sceptical about reasons, and implies that it is false that we have reason to be sceptical about reasons (2002: xxv).

Terence Cuneo makes a similar argument against a proponent of 'epistemic nihilism', roughly the epistemic counterpart to Mackie's error theory – the view that there are no epistemic facts, that there are no epistemic reasons for belief, and so on:

---

[20] For reasons of space, I cannot possibly hope to outline *all* of the arguments for the epistemic existence premise. I focus, therefore, on just two: *the argument from self-defeat*, to be outlined presently, and two instances of *the argument from epistemic merits and demerits*, to be outlined below.

> Either epistemic nihilists hold that we have reasons to believe epistemic nihilism or that we do not. If epistemic nihilists hold that we do have reasons to believe their position, then their position is self-defeating in the sense that it presupposes the existence of the very sorts of entity that it claims do not exist. After all, if we have reasons to believe epistemic nihilism, then it follows by conceptual necessity that we rationally ought, all other things being equal, to believe it. But there are no rational oughts according to epistemic nihilists; there are no facts that imply that certain propositions are belief-worthy or that failing to believe something on good available evidence renders one (all other things being equal) irrational (2007: 117).

The idea is that the epistemic existence premise is true because one cannot coherently deny it: if you deny that there are any epistemic reasons, for example, then you cannot coherently believe that there are no epistemic reasons, since you have epistemic reasons to believe that there are no epistemic reasons.[21] Cowie (2014: 411) reconstructs this general line of argument as follows:

(1) If the epistemic error theory is true, then there are no epistemic reasons for belief.

(2) If there are no epistemic reasons for belief, then there is no epistemic reason to believe the epistemic error theory.

(3) So, either the epistemic error theory is false or there is no epistemic reason to believe it.

(4) Therefore, we should reject the epistemic error theory.

(5) Therefore, the epistemic existence premise is true.

Cowie argues that the argument from self-defeat – if sound – has the upshot that we should reject the parity premise. For if the argument from self-defeat is successful, then it shows only that there is a sufficient reason for rejecting the

---

[21] Granted, it doesn't follow, from the fact that we have no reason to believe a given theory, that the theory is false (cf. Olson 2011; Streumer 2017). But we can assume for the sake of argument – as Cowie does (2014: 411) – that this reply to the argument from self-defeat misses the mark. Cowie wants to argue that, even if we concede that the argument from self-defeat warrants acceptance of the epistemic existence premise, the argument still succumbs to the objection from disparity.

epistemic error theory. But this, according to Cowie, undermines the parity premise:

> [A]n analogue of [the argument from self-defeat] certainly would not show there to be a sufficient reason for rejecting the moral error theory. This is because the moral error theory would have the consequence that there are no moral reasons. But it would not have the consequence that there are no epistemic reasons. And so, the truth of the moral error theory would be compatible with the existence of a reason to believe that theory. This disanalogy undermines the parity premise. It undermines the parity premise because it entails that there is a sufficient reason for rejecting the epistemic error theory (namely, that it is self-defeating), that is not also a sufficient reason for rejecting the moral error theory (2014: 411).

This, according to Cowie, is one instance of the objection from disparity. Consider another argument for the epistemic existence premise: *the argument from global scepticism* (Rowland 2013; cf. Cuneo 2007: 119). According to the argument from global scepticism, the epistemic existence premise must be true, because if it were false, it would imply global scepticism. Essentially, the idea is that if there are no epistemic reasons (i.e. if the epistemic existence premise is false), then *nobody knows anything*. Since we do know some things, the epistemic existence premise must be true.

The argument from global scepticism appeals to a connection between reasons for belief and epistemic justification. As Rowland (2013: 13) points out, the following claim is relatively uncontroversial in epistemology:

(A) If *S* knows that *p*, then there is some epistemic justification for believing that *p*.

We can only know that it is raining outside if we have some justification for believing that it is raining outside; I cannot know that snow is white unless I have justification for believing that snow is white; and so on. We can combine (A) with the following claim:

(B) If there is some epistemic justification for believing that p, then there is an epistemic reason for believing that p.

Claim (B), too, seems plausible. Rowland writes:

> It is deeply plausible that for p to justify q just is for p to be a reason for q, and that epistemic justifications just are epistemic reasons, or that epistemic reasons are just epistemic justifications. But regardless, it seems that if there is justification for a belief, then there is a reason to believe it. It would be extremely odd to claim that I am justified in believing that dinosaurs once roamed the earth but there is no reason at all for anyone to believe this (2013: 13).

So it looks like (B) is plausible. If both (A) and (B) are plausible, then we have an argument for the epistemic existence premise. For when we combine (A) and (B), we get the following hypothetical syllogism:

(A) If S knows that p, then there is some epistemic justification for believing that p.

(B) If there is some epistemic justification for believing that p, then there is an epistemic reason for believing that p.

(C) Therefore, if S knows that p, then there is an epistemic reason for believing that p.

That is, if there is an epistemic reason to believe that p, then the epistemic error theory is false, since the epistemic error theory would say that there are no epistemic reasons for belief. The argument from global scepticism – if sound – constitutes a defense of the epistemic existence premise because it entails a highly unpalatable consequence: the consequence that, if the epistemic error theory is true, then we have no knowledge. For simplicity's sake, Cowie (2014: 413) reconstructs Rowland's argument from global scepticism as follows:

(1) If the epistemic error theory is true, then there are no epistemic reasons for belief.

(2) If there are no epistemic reasons for belief, then we are never justified in believing a proposition.

(3) If we are never justified in believing a proposition, then we have no knowledge.

(4) Therefore, if the epistemic error theory is true, then we have no knowledge (1, 2, 3).

(5) But we do have some knowledge.

(6) Therefore, the epistemic error theory is false (4, 5).

(7) Therefore, the epistemic existence premise is true.

Cowie suggests, however, that the argument from global scepticism is really an instance of a more general argument, due primarily to Cuneo (2007: 119). It is worth briefly outlining this more general argument, since this is the argument Cowie attacks – and Cowie's objections to the more general argument will, if successful, also undermine the argument from global scepticism. According to the more general argument, "the truth of the epistemic error theory entails the lack of any *epistemic merits or demerits*; any facts about justification, warrant, rationality, knowledge, epistemic reasons, and so forth" (Cowie 2014: 413). Here is Cowie's reconstruction of the more general argument:

(1) If the epistemic error theory is true, then there are no normative epistemic facts.

(2) If there are no normative epistemic facts, then there are no epistemic merits or demerits.

(3) Therefore, if the epistemic error theory is true, then there are no epistemic merits or demerits (1, 2).

(4) But there are some epistemic merits or demerits.

(5) Therefore, the epistemic error theory is false.

(6) Therefore, the epistemic existence premise is true.

Call this more general argument *the argument from epistemic merits and demerits* (Cowie 2014: 414). Now, the objection from disparity will *not* work against the argument from epistemic merits and demerits. Cowie writes:

Suppose that the argument from epistemic merits and demerits is sufficient to warrant rejection of the epistemic error theory. What would follow? One might think that the parity premise would be undermined in another instance of the (now familiar) objection from disparity. In this case, however, the argument is too quick. The objection from disparity doesn't work. This is because, plausibly, the argument from epistemic merits and demerits does have a moral analogue that warrants rejection of the moral error theory. Specifically, the lack of moral merits and demerits entailed by the moral error theory may be a sufficiently undesirable result to warrant rejection of that theory. If there are no moral merits or demerits, then there is nothing morally non-meritorious about, for example, torturing babies merely for fun. But, this is clearly false, and warrants rejection of the moral error theory (2014: 414).

The idea is that the objection from disparity does not work against the argument from epistemic merits and demerits, because unlike the argument from self-defeat, the argument from epistemic merits and demerits *does* have a moral analogue: the argument from *moral* merits and demerits.

However, Cowie argues that the immunity of the argument from epistemic merits and demerits to the objection from disparity leaves it open to another: the objection from redundancy. That is, while the argument from epistemic merits and demerits avoids the first horn of Cowie's dilemma (a), it lands upon the second (b). According to Cowie, the immunity of the argument from epistemic merits and demerits to the objection from disparity is premised on the following claim:

(P) The argument from epistemic merits and demerits has a moral analogue (the argument from moral merits and demerits) that warrants rejection of the moral error theory (2014: 415).

Cowie argues that if (P) is true, then the argument from epistemic merits and demerits falls into the second horn of the dilemma (b): the companions in guilt argument becomes dialectically redundant. For if (P) is true, then there is a *direct* argument against the moral error theory (where a 'direct argument' is one that doesn't proceed by analogy), which might run as follows:

27

(1) If the moral error theory is true, then there are no moral merits or demerits.

(2) There are moral merits and demerits.

(3) Therefore, the moral error theory is false.[22]

But if this argument gives us sufficient *direct* reason to reject the moral error theory, then the CG argument seems to be rendered dialectically redundant. Cowie writes:

> [The companions in guilt argument] becomes dialectically redundant in that one would have a reason to accept the parity premise, and hence, the companions in guilt argument, only if one already possesses a sufficient, direct, argument for the falsity of the moral error theory. The companions in guilt argument would, then, be an idling wheel. One would be rationally required to accept its conclusion that the moral error theory is false only if one already took oneself to have a sufficient reason to accept that the moral error theory is false (2014: 415).

Cowie acknowledges, however, that there is a sense is which the CG argument is *not* dialectically redundant: it is not dialectically redundant in the sense that it would (if sound) provide *further* reasons for rejecting the moral error theory, for a philosopher who *already* rejects the moral error theory. Cowie concedes this, but notes that the CG argument *is* dialectically redundant *relative to the aim of the argument*. The aim of the argument is to provide reasons for rejecting the moral error theory for someone who does *not* already reject the error theory. This is the sense in which Cowie claims the CG argument is rendered dialectically redundant.

---

[22] Granted, Mackie is not an eliminativist: he does not advocate for the elimination of moral concepts and moral discourse. Rather, he thinks we follow subsidiary norms based on a particular conception of human flourishing, and cashes out the *point* of moral discourse in terms of these subsidiary norms. Read most charitably, however, Cowie's point is this: 'epistemic merits and demerits' includes *facts* about knowledge, justification, and so on; by analogy, therefore, 'moral merits and demerits' would include *moral facts*. Since Mackie clearly doesn't think there are moral facts, it is plausible that the argument from epistemic merits and demerits does have a moral analogue.

So, if Cowie is right, then companions in guilt theorists are in trouble: if they appeal to the argument from self-defeat in order to vindicate the epistemic existence premise, then they seem to end up undermining the parity premise. They succumb to the first horn of Cowie's dilemma (a): the objection from disparity. But if, on the other hand, they appeal to an argument which does not succumb to the first horn of the dilemma – such as the argument from global scepticism or the argument from epistemic merits and demerits – then they fall into the second horn of the dilemma (b): the CG argument is rendered dialectically redundant. Cowie suggests that a general (if ambitious) moral can be drawn from the foregoing story: it is not just that a few particular companions in guilt arguments *do not* work, but that companions in guilt arguments – as such – *cannot* work, for "[a]ny instance of the companions in guilt argument will be undermined by either the objection from disparity or the objection from redundancy" (2014: 415).[23]

But is Cowie right? In response to Cowie's criticisms, defenders of CG arguments have either attempted to show how, contra Cowie, CG arguments *can* work (Rowland 2016), or they have attacked Cowie's own arguments (Das 2016). Let us look at just the first of these responses.[24] Rowland (2016: 161) wants to defend the following argument[25] against the objections Cowie raises:

(i) According to the moral error theory, moral reasons are metaphysically problematic because they are categorically normative.

---

[23] In order to further support the objection from disparity and the objection from redundancy, Cowie (2014: 416-21) draws an analogy with the philosophy of mathematics – in particular, the Putnam-Quine 'indispensability thesis', which concerns the indispensability of mathematical objects to natural science. While Cowie's discussion is highly interesting, to outline it would be to take us too far afield. See Rowland (2016: 169-70) for a response.

[24] In more recent work, Cowie (2016) has attempted to develop a "master argument" against CG strategies. I cannot go into the details of this argument here, both for reasons of space and because I think it will take us too far afield. For a reply, see Das (2017), who argues that Cowie's master argument fails and that there is little hope for any such forthcoming master argument.

[25] Note that this argument is formulated somewhat differently to Rowland's original argument from epistemic reasons. It seems to be a synthesis of the companions in guilt argument as presented in Cuneo (2007), Cowie (2014), and Rowland (2013). Nothing should turn on this, since Rowland reads Cowie as disputing only the move from (iv) to (v) in this argument.

(ii) But epistemic reasons are categorically normative.

(iii) And there are epistemic reasons.

(iv) Therefore, there are categorically normative properties.

(v) Therefore, moral error theorists' scepticism about moral reasons, on the grounds of moral reasons' categorical normativity, is unwarranted.

According to Rowland, the difficulties Cowie raises for CG arguments concern the move from the subconclusion (iv) to the main conclusion (v), because Cowie accepts (for the sake of argument) that epistemic reasons are categorically normative and that they exist:

> The idea here is that the move from [iv] to [v] is not justified because to show that there are some categorically normative properties is not to show that categorically normative properties are not metaphysically problematic. And, so long as categorically normative properties are metaphysically problematic, showing that there are some facts E that are categorically normative will not show that there are no grounds for scepticism about facts M on the grounds of M's categorical normativity (Rowland 2016: 161-2).

First, according to Rowland, we should hold that "*either* epistemic reasons and moral reasons are instances of fundamentally the same relation with different relata, *or* epistemic and moral reasons are metaphysically unproblematic" (2016: 162). And if moral reasons and epistemic reasons are instances of fundamentally the same relation with different relata, then according to Rowland, the main conclusion (v) follows (assuming there are epistemic reasons). Let us flesh this out.

To begin with, recall the initial worry raised for the idea that epistemic reasons and moral reasons stand or fall together: the former concern reasons for *belief*, whereas the latter concern reasons for *action*. Rowland's way of dealing with this issue was to point out that both reasons for action and reasons for belief can be thought of as *different relata* of the same normative relation: that of *being a reason for*. (Again, reasons' being relations does not mean they are not also facts.) As Stratton-Lake puts the point:

> The difference between practical and epistemic reasons is not that they stand in different warranting relations to certain things (one relation queer and the other innocuous), but that they warrant different things. Practical reasons warrant pro-attitudes and actions, whereas epistemic reasons warrant beliefs (2002: xxv-vi).

According to Rowland, we have good reason to accept that moral and epistemic reasons are instances of the same fundamental relation (2016: 163). For one thing, both proponents of CG arguments *and* Cowie accept that epistemic reasons are categorical. (Cowie accepts this for the sake of argument.) Secondly, according to Rowland:

> [I]f we hold that epistemic and moral reasons are fundamentally different relations, then 'reason' in 'that there are dinosaur bones is a reason to believe dinosaurs once roamed the earth' has as much in common with 'reason' in 'that she'll die if you don't save her is a reason to save her' as 'bank' in 'that is a river bank' has in common with 'bank' in 'that is a financial bank', that is absolutely nothing beyond sounding the same when pronounced. But 'reason' in 'that there are dinosaur bones is a reason to believe dinosaurs once roamed the earth' and 'reason' in 'that she'll die if you don't save her is a reason to save her' do not have nothing in common (2016: 163).

Now, either moral and epistemic reasons are instances of the same fundamental relation in the way just described, *or* they are metaphysically unproblematic. For we might think that moral and epistemic reasons are *not* of the same normative relation if it were easier to reduce epistemic reasons to natural or non-normative facts than it is to reduce moral reasons to natural or non-normative facts (Rowland 2016: 164). But Cowie accepts for the sake of argument, and proponents of CG arguments have argued at length, that it is not easier to reduce epistemic reasons to natural or non-normative facts in this way (cf. Cuneo 2007; Rowland 2013; Cowie 2014: 410). Thus according to Rowland, "if both epistemic and moral reasons can be reduced to natural/non-normative facts, moral reasons are not metaphysically strange since they are identical to natural/non-normative facts" (2016: 164).

Rowland now argues that main conclusion of the above CG argument (v) follows from (iv) if moral and epistemic reasons are of the same fundamental relation. Suppose it is not obvious that moral and epistemic reasons are not metaphysically problematic, and so we can say that they are instances of the same fundamental normative relation. Let *relation R* be the fundamental normative relation of which moral and epistemic reasons are instances. According to Cowie (2014: 416), proponents of CG arguments show that there are some cases of categorical reasons that have certain *special* properties, such as the property of being self-defeating (Cuneo 2007) or the property of being entailed by knowledge (Rowland 2013). But Cowie argues that to show that there are some cases of categorical reasons which have certain special properties only undermines scepticism about the particular categorical reasons that have these properties. It does not undermine scepticism about other instances of categorical reasons – such as moral reasons. But according to Rowland:

> [I]f epistemic and moral reasons are both fundamentally Rs, and there are instances of relation R – as Cowie assumes – then the only thing that could metaphysically count as positing more instances of R on the bases of Rs' categorical normativity is a requirement of *quantitative* parsimony. That is, a requirement according to which it is better not to *posit more of the same kind of thing*, rather than a requirement of *qualitative* parsimony according to which it is better not to *posit new kinds of things*. (Since, to hold that there are more Rs is not to hold that there is a new kind of thing.) (2016: 165).

To illustrate this, we can consider two theories:

(THEORY A): There are categorical reasons beyond epistemic reasons.

(THEORY B): The only categorical reasons that exist are epistemic reasons.

According to Rowland, if quantitative parsimony "is a consideration that can be outweighed by *any competing consideration*, then the fact that [THEORY B] is more quantitatively parsimonious than [THEORY A] will be outweighed by the fact that [THEORY B] entails that much of our discourse and practice is in vast error; since we believe that there are moral reasons to do things" (2016: 165).

However, according to Rowland it is difficult to see how such a quantitatively parsimonious theory would provide a better explanation of anything, when it comes to epistemic reasons. Indeed, since such a quantitatively parsimonious theory would not be able to able to explain how much of our discourse is not in error, it seems *less* explanatory when it comes to epistemic reasons. Thus Rowland maintains that without further justification it is difficult to see why we should hold that there are as few categorical reasons of the same fundamental sort as possible:

> It might seem that Cowie's point is that proponents of companions in guilt arguments such as Cuneo do not show that there are instances of relation *R tout court*. But rather only show that there are particular kinds of instances of relation *R*, namely instances of *R* that have further 'special' properties such as the property of being self-defeating to deny the existence of (Cuneo) or the property of being entailed by knowledge that there must be (Rowland). But it could only be quantitative parsimony that would count against positing such instances of *R* that do not have these properties. Since to hold that there are instances of *R* that do not have these 'special' properties is not to posit a new type of relation, but merely to posit more instances of a relation that there must be, that we should already be committed to there being instances of, or that it is self-defeating to not be committed to. And, it is not plausible that an instance of *R* that is entailed by knowledge that there must be or is an instance of *R* that it is self-defeating to deny the existence of is a fundamentally distinct relation from instances of *R* that do not have such properties (Rowland 2016: 166).

Of course, more needs to be said, but we at least have a *prima facie* case against Cowie's objections to CG arguments: if moral reasons and epistemic reasons are instances of the same fundamental normative relation, then CG arguments *can* work. And that is all we need for our present purposes.

1.4.2 DOXASTIC NORMATIVISM AND COMPANIONS IN GUILT

We have seen, then, that the first issue – the issue of whether there is sufficient parity between moral and epistemic facts in order for a CG argument to work – does not undermine CG strategies: contra Cowie, there is no internal tension

between the parity premise and the epistemic existence premise. It seems plausible that something similar works in the case of doxastic normativity. But what of the second issue, the issue of whether there are categorical doxastic norms? And if there are categorical doxastic norms, what grounds them? It is to this issue that we now turn.

In recent years, epistemology has taken a *normative* turn. As in ethics, in epistemology there are *epistemic deontologists* (e.g. Boghossian 2003), *virtue theorists* (e.g. Zagzebski 2003), and *consequentialists* (e.g. Petersen 2013). Naturally, CG theorists appeal to epistemic deontology, since on this view, there are categorical norms according to which you ought to believe in accordance with your evidence, for example, or you ought to believe what it is rational to believe.[26] I want to develop a CG argument based on a certain metaphysical thesis: *doxastic normativism*. Doxastic normativism is the thesis that norms are constitutive of, or essential to, belief. Belief, some say, has a mind-world 'direction of fit' (cf. Anscombe 1957: 56 ff.); it 'aims at truth' (cf. Williams 1973: 136-7). There are two issues which arise from the slogan 'belief aims at truth', one metaphysical and one epistemological (Hattiangadi 2010). The epistemological issue is about the epistemic evaluation of belief – it is about whether truth, knowledge, or whatever, is of fundamental epistemic value. The metaphysical issue, however, is about the essential nature of belief, and this is the issue with which we shall primarily be concerned. Doxastic normativists interpret the slogan 'belief aims

---

[26] One might worry that there is a more fundamental issue here. Doesn't the idea that there are norms on belief presuppose a strong form of *doxastic voluntarism*, the view that we have voluntary control over our beliefs? Indeed, in his classic presentation of the issue, Bernard Williams (1973) invokes the slogan "belief aims at truth" in order to *explain* why we cannot believe at will. It might be objected, therefore, that before CG arguments can even get off the ground, we need to deal with the question of doxastic voluntarism. I cannot do this here. As such, I will assume that we have *some* level of voluntary control over our beliefs. However, it is worth noting that a number of philosophers have argued explicitly against the idea that belief's having an aim, or there being norms on belief, presupposes doxastic voluntarism. McHugh (2012a), for example, argues that doxastic voluntarism is false, but that this does not undermine epistemic deontology. And Reisner (2013) argues that belief's having an aim – whether that aim is truth or knowledge – is not in tension with doxastic voluntarism. See also Owens (2000) and Chrisman (2008) for relevant discussion.

at truth' metaphorically.[27] After all, beliefs are not agents, and so do not literally aim at anything; and nor does it seem to be *believers* that do the aiming, since believers typically do not form their beliefs with the conscious aim of getting at the truth. *Teleologists* about the aim of belief, on the other hand, interpret 'belief aims at truth' literally: either believers typically aim at truth when forming their beliefs, or natural selection has endowed them with truth-conducive belief-forming mechanisms (Velleman 2000; Steglich-Petersen 2006). Teleologists want to give a naturalistic account of the aim of belief, normativists a normative account.

According to the doxastic normativist, categorical norms are in some sense *constitutive* of belief. [28] Doxastic normativists typically accept (some version of) the *truth norm*:

(TRUTH): For any subject S and proposition p, S ought to believe that p if and only if p is true.[29]

The teleologist, on the other hand, will accept the truth norm, only deny that it is categorical (Steglich-Petersen 2010: 749). Rather, the norm is *instrumental* as a means of getting at the truth: essentially, beliefs require having an aim or

---

[27] See, e.g., Wedgwood (2002), Boghossian (2003), Shah (2003), Shah and Velleman (2005), Engel (2007), and Whiting (2010).

[28] Although in Chapter 2 I will say more about what constitutivity in this context amounts to, the notion of constitutivity or essentiality we are working with remains imprecise and somewhat unclear. According to Glüer and Wikforss (2013: 81), constitutivity or essentiality in this context need not amount to anything more than metaphysical necessity. Moreover, some philosophers hold that the relevant notion of necessity is *conceptual* necessity (Boghossian 2003; Shah 2003; Shah and Velleman 2005). But we can assume for the sake of argument that conceptual necessity *entails* metaphysical necessity. Glüer and Wikforss note that while this assumption is not entirely uncontroversial, it is harmless in this context, since it is endorsed by those normativists (*op. cit.*) who construe normativism as a conceptual truth. The idea is that we want to say something about the *nature* of belief, not the *concept* of belief. But if conceptual necessity entails metaphysical necessity, then our considerations of the concept of belief will apply *mutatis mutandis* to the nature of belief.

[29] Note two things. First, (TRUTH) seems obviously too strong, but let us assume it for the moment. I discuss the formulation of doxastic norms in Chapter 2. Second, I argue in Chapters 2 and 3 that *knowledge*, rather than merely truth, is the aim and norm of belief. Again, let us postpone discussion of this in order to keep matters simple.

intention to believe only the truth, and norms like (TRUTH) are instrumental as a means to this end.

But before saying more about normativism, and how it bears on CG arguments, it is worth giving some context, since the context will become especially relevant later on. Doxastic normativism emerges as one strand of the so-called 'rule-following considerations' (cf. Boghossian 1989). Since the publication of Saul Kripke's *Wittgenstein on Rules and Private Language* (1982) – an influential reading of Ludwig Wittgenstein's *Philosophical Investigations* (1953) – much discussion has been generated over the alleged normativity of *linguistic meaning*, though many take the normativity thesis to apply also to mental content.[30] Kripke's Wittgenstein (KW) presents a 'sceptical paradox' according to which there are no facts in virtue of which speakers *mean* anything by their linguistic expressions; there is no fact which constitutes, or makes it the case that, a speaker means one thing rather than another by a given expression. To use Kripke's example, suppose a bizarre sceptic queries an interlocutor as to what *makes it the case* that he means *addition*, rather than *quaddition*, by 'plus' and '+', where quaddition (symbolized by '⊕') is defined as follows:

$$x \oplus y = x + y, \text{ if } x, y < 57$$

$$= 5 \text{ otherwise}$$

If our interlocutor means *addition* by 'plus', then when queried about the answer to the question '68 + 57 = ?" he should answer with the *sum*: '125'. If, on the other hand, he means *quaddition*, then he should answer with the *quum*: '5'. KW's sceptical challenge is to point to some *fact* – about my mental state or behavioral dispositions – which constitutes, or makes it the case that, I mean *addition* rather than *quaddition* by 'plus', *green* rather than *grue* by 'green', and so on.

---

[30] Arguably, something like the normativity thesis can be traced back to Kant's *Critique of Pure Reason* (1781). See Brandom (1994) for such a reading of Kant.

Essentially, Kripke argues by a process of elimination: he considers all the sorts of fact which purport to constitute meaning and finds none of them adequate. Among those facts which purport to constitute meaning are facts about *dispositions*. According to semantic dispositionalism (at least in its crudest form), facts about speakers' dispositions constitute what they mean: I mean *addition* by 'plus' and '+' if I am disposed to use 'plus' to answer addition-queries with the *sum* of the relevant numbers, rather than to answer according to some other function. But Kripke argues, among other things, that one of the ways in which dispositionalism falters is in its failure to account for the *normativity* of meaning:

> Suppose I do mean addition by '+'. What is the relation of this supposition to the question how I will respond to the problem '68 + 57'? The dispositionalist gives a *descriptive* account of this relation: if '+' meant addition, then I will answer '125'. But this is not the proper account of the relation, which is *normative*, not descriptive. The point is *not* that, if I meant addition by '+', I *will* answer '125', but that, if I intend to accord with my past meaning of '+', I *should* answer '125' […] The relation of meaning and intention to future action is *normative*, not *descriptive* (1982: 37).

The idea is that if a speaker is to accord with her past meaning of 'plus' in the present, then she *ought* to answer addition-queries with their sum; in particular, she should respond to the above query with '125'. This seems to suggest that there is something essentially normative about meaning, for we can derive *oughts* from meanings: if I mean addition by 'plus', then I *ought* to answer with the sum. Of course, the example need not be a mathematical one: if I mean *green* by 'green', then I ought to apply 'green' only to green things.

More recently, however, Kripke's normativity thesis in the semantic case has come under scrutiny.[31] Instead, the focus has shifted to *mental content* (Glüer and Wikforss 2009: 31-2). Just as some claim that norms are essential to meaning, some claim that norms are essential to content. Doxastic normativism is one

---

[31] See, e.g., Glüer (1999), Glüer and Pagin (1999), Wikforss (2001), and Hattiangadi (2007). See Hlobil (2015) for a response to some of these criticisms.

strand of content normativism, though either thesis can be endorsed quite independently of the other: one can be a doxastic normativist without being a content normativist, and vice versa. (Granted, arguments for content normativism often go hand-in-hand with those for doxastic normativism – for instance, some philosophers, such as Boghossian (2003, 2005), argue that the normativity of content can be defended *via* the normativity of belief.)

Doxastic normativism ('normativism' hereafter) if true provides a neat picture of why we ought to follow the epistemic norms, such as norms of evidence or rationality. As we shall see, normativists such as Boghossian (2003, 2005) hold that the 'objective' norm of truth *grounds* the 'subjective' epistemic norms and supplies their 'rationale' (more on this in Chapter 2). The epistemic norms hold *because* the truth norm does. If normativism is true, then, we have an account of the epistemic norms. Moreover, one of the motivations for normativism is that the normativity of belief purports to distinguish belief from the other cognitive attitudes, such as imagining, wishing, or hoping. The idea is that belief is truth-directed, but involves norms: when you imagine that p, you imagine that p is true, but there are no *oughts* constraining what you are permitted to imagine. When you believe that p, you believe that p is true, but you *ought* to believe that p iff p. (Again, more on this in Chapter 2.)

How is all this relevant to CG arguments? According to normativism, categorical norms such as (TRUTH) are *constitutive* of belief; that is, part of *what it is* to be a belief or a believer is to be governed by these norms. Nothing not governed by these norms is a belief. Extant CG arguments conclude that the argument from queerness is insufficient, because there *are* categorical epistemic norms, and the error theorist relies on the premise that there are no such norms. But if normativism is true, then the conclusion of a CG argument is far more radical than that: a doxastic analogue of the argument from queerness can be run in the belief case, which has the paradoxical upshot that *there are no beliefs and no believers*. For the normativist holds that categorical norms are constitutive of

belief, and that this is a conceptual truth; but if the argument from queerness is sound, then there is nothing in the world which instantiates our concept of a doxastic fact (i.e., there are no categorical norms which constitute belief). Doxastic judgements (for example: the belief "Sarah believes that snow is white") would be systematically and uniformly false.[32]

This, then, is the threat to the error theory: assuming normativism, the argument from queerness (if sound) can be run in the doxastic case, and so would threaten to support an error theoretic conclusion about belief. This is bad, and so I take it that this constitutes something of a *reductio* against the argument from queerness. In the next section, I will flesh this out.

1.5 COGNITIVE SUICIDE

How might normativism allow for a doxastic analogue of the argument from queerness? A doxastic analogue of the argument from queerness would threaten to undermine what we might call *doxastic realism*. Let doxastic realism be the following view:

> (DOXASTIC REALISM): Doxastic judgements (i) purport to state facts; and (ii) some doxastic judgements are true.

According to (DOXASTIC REALISM), in making a doxastic judgement (e.g. "Sarah believes that snow is white") the judgement purports to truly attribute a belief to the agent, and the belief attribution is true just in case the agent does hold that belief. When, for instance, I judge that Sarah believes that snow is white, my judgement expresses a second-order belief to the effect that Sarah believes that snow is white, and it has a true content if and only if Sarah does believe

---

[32] This thesis thus seeks to complement the discussion of Evans and Shah (2012). According to them, the 'normativity of the attitudes' – which includes belief as such a normative attitude – greatly limits the number of options in metaethics. In particular, Evans and Shah argue that (their version of) normativism is incompatible with metaethical anti-realism: "none of the traditional anti-realist metaethical views can fully accommodate certain central features of mental agency – features that every human thinker is closely familiar with" (2012: 80). I will not go into the details of this argument here.

that snow is white. (DOXASTIC REALISM) is thus a straightforward analogue of (MORAL REALISM).

Now, recall that Mackie asks both a *conceptual* question and an *ontological* question about the nature of putative moral facts. Likewise, in order to develop a doxastic analogue of the argument from queerness, we need to ask, first, what our concept of a doxastic fact is, and second, whether there is anything in the world which instantiates that concept. Now, many normativists construe normativism as a *conceptual truth* about belief (Boghossian 2003; Shah 2003; Shah and Velleman 2005). On this view, (TRUTH), or something like it, is part of the very concept of belief; it is constitutive of belief.[33] And if this is right, we can imagine a doxastic analogue of moral rationalism which gets the argument from queerness off the ground.[34] It might read as follows:

> (DOXASTIC RATIONALISM): It is a conceptual truth that if S believes that p,
> then S has a reason to believe that p if and only if p is true.

Thus according to (DOXASTIC RATIONALISM), it is a conceptual truth that if Sarah believes, for instance, that snow is white, then she has a reason to believe that snow is white iff snow is white. As in the moral case, we can grant the doxastic error theorist her commitment to (DOXASTIC RATIONALISM). If we grant this, then it is plausible that our concept of a doxastic fact is the concept of a reason for belief, i.e. a reason for believing only the truth. But is it a *categorical* reason for belief – that is, an *objectively prescriptive* feature of the world? Assuming

---

[33] Again, I will assume for the sake of argument in this chapter that (TRUTH) is the constitutive norm of belief. I argue in Chapters 2 and 3 that *knowledge* is the aim and norm of belief. However, this shouldn't affect our present considerations: I will argue that, assuming normativism, a doxastic analogue of the argument from queerness can be developed, regardless of whether the normativity in question is construed along the lines of the truth norm or along the lines of the knowledge norm.

[34] See Mitova (2011) for a highly interesting discussion of classic metaethical debates transposed onto epistemology. One such debate is the debate over motivational internalism: Mitova fleshes out a position she calls *epistemic internalism*, which should be complementary to the arguments of this Chapter. It seems plausible that, just as in the moral case, (DOXASTIC RATIONALISM) will entail some version of what we might call *doxastic internalism*, but we need not flesh this out for our present purposes.

normativism, then it is plausible that this is our concept of a doxastic fact. For if normativism is true, then it is a conceptual truth that belief is norm-governed. Given this assumption, then, we might reconstruct a doxastic analogue of the argument from queerness as follows:

(1*) Our concept of a doxastic fact is a concept of a reason for belief.

(2*) If our concept of a doxastic fact is a concept of a reason for belief, then our concept of a doxastic fact is a concept of an objectively prescriptive feature of the world (i.e., a *categorical* reason for belief).

(3*) Therefore, our concept of a doxastic fact is a concept of an objectively prescriptive feature of the world.

(4*) But if our concept of a doxastic fact is a concept of an objectively prescriptive feature of the world and there are no objectively prescriptive features of the world, then doxastic judgements are systematically and uniformly false.

(5*) There are no objectively prescriptive features of the world.

(6*) Therefore, doxastic judgements are systematically and uniformly false.

The first hurdle this argument faces is in defending premise (2*). Whether (2*) is true seems to depend on whether *normativism* is true, since normativism is the view that it is part of our concept of belief that belief is norm-governed and that such norms are constitutive of belief. However, let us grant (2*) for the moment. And if we grant this, then subconclusion (3*) follows straightforwardly from (1*) and (2*). Premise (4*) seems obvious: just as our unicorn judgements are systematically and uniformly false because there are no unicorns, our doxastic judgements (e.g. the judgement "Sarah believes that snow is white") would be systematically and uniformly false if there were nothing in the world that instantiates our concept of a doxastic fact, i.e. if there were no objectively prescriptive features of the world or categorical reasons for belief. As for (5*), this is just the same as premise (5) in Mackie's argument from queerness: the idea is that *if* there were any objectively prescriptive features of the world, they would

be metaphysically and epistemologically queer. And the main conclusion (6*) follows from (3*), (4*), and (5*).

Now, suppose this doxastic analogue of the argument from queerness is sound. What would the upshot be? I take it that – given normativism and our foregoing discussion of the parity premise in CG arguments – the conclusion of the doxastic argument from queerness is sufficiently unpalatable to warrant rejection of the moral error theory. For the conclusion of the doxastic argument from queerness is an error theoretic one: doxastic judgements are systematically and uniformly false. This has the consequence that there are no doxastic facts – that there are *no beliefs and no believers*. For if doxastic judgements or belief ascriptions are always false, then it is false that:

I believe that it rained today.

Sarah believes that snow is white.

Marco believes that Mallory reached the summit.

Bilbo believes he hasn't had his second breakfast.

And so on. If it is false that anyone has ever believed anything, then it seems to follow that there are no doxastic facts – just as it follows, from the conclusion that *moral* judgements are systematically and uniformly false, that there are no moral facts. In other words, (DOXASTIC REALISM) would be false.

But why is the conclusion that there are no doxastic facts (i.e., that there are no beliefs and no believers) really so unpalatable? For one thing, it is highly counterintuitive. But this does not mean the doxastic error theory is false. However, the conclusion that there are no beliefs and no believers would amount to a denial of the *common sense conception of the mental* – the conception according to which there are beliefs, desires, and other mental states familiar to folk psychology. According to Lynn Rudder Baker (though in the context of physicalism),

such a denial would amount to a kind of *cognitive suicide*: "Since cognition without content is empty, denial of the common sense conception of the mental may be a kind of cognitive suicide that we are constitutionally unable to commit to" (1998: 18). We are constitutionally unable to commit to the doxastic error theory because its upshot is the paradoxical one that there are no beliefs and no believers.[35] This upshot is paradoxical because it is logically impossible to truly believe that there are no believers. For if you believe that there are no believers, then it is false that there are no believers; and if it is true that there are no believers, then you don't believe that there are no believers.

The upshot of the supposition that our doxastic analogue of the argument from queerness is sound, then, can be represented with the following argument – an argument which purports to show that, assuming normativism, the proper conclusion of a CG argument is far more radical than its proponents take it to be:

(1) If normativism is true, then it is part of the concept of belief that it is constituted by categorical norms.

(2) Normativism is true (or so we are supposing).

(3) Therefore, it is part of the concept of belief that it is constituted by categorical norms (1, 2).

(4) But if the argument from queerness is sound, then there are no categorically normative features of the world.

(5) Therefore, if the argument from queerness is sound, then there is nothing in the world instantiating our concept of belief (3, 4).

(6) Therefore, if the argument from queerness is sound, then there are no beliefs and no believers.

The conclusion of many extant CG arguments, such as that of Rowland (2013), is that moral error theorists' scepticism about moral reasons is unwarranted,

---

[35] Indeed, "There are no believers" is a standard example of a Moore-paradoxical proposition, or 'blindspot' (Sorensen 1988). I discuss blindspots in Chapter 2.

because there are categorical epistemic reasons: if there were no categorical epistemic reasons, then judgements of the form "The belief that *p* is justified" would be systematically and uniformly false. But if what I have said above is correct, then assuming normativism, the conclusion of a CG argument is far more radical than that: it is not that moral error theorists' scepticism about moral reasons on the grounds of their categorical normativity is unwarranted, but that the argument from queerness must be unsound because it has the paradoxical consequence that there are no beliefs and no believers. Belief ascriptions of the form "*S* believes that *p*" would be systematically and uniformly false. If this is right, then this consequence, I take it, constitutes a *reductio ad absurdum* against the argument from queerness.

In response to CG strategies, we have seen that some philosophers attack either the parity premise or the idea that there are categorical epistemic norms. But Jonas Olson (2011) – one of the few philosophers in the debate over CG arguments who explicitly links these arguments to normativism – accepts for the sake of argument that the CG strategy goes through and assesses its implications. According to him, the implications of *epistemic* error theory are not so unpalatable as to warrant rejection of the moral error theory:

> Since epistemic error theory holds that no claims about what there is reason for agents to believe and not to believe are true, normativism seems to imply that according to epistemic error theory no belief ascriptions are true […] The simplest and most plausible response on behalf of the epistemic error theorist is to agree with the normativist that [norms of evidence, rationality, etc.] are distinctive of belief and belief ascriptions. But the error theorist should insist that these are immanent [i.e., hypothetical] and not transcendent [i.e., not categorical] norms. In order to distinguish belief from other attitudes such as conjecture, imagining, supposition, wishful thinking, and the like, it is not necessary to

assume that the norms associated with belief are transcendent; it suffices to assume that they are immanent (Olson 2011: 92).[36]

In essence, the idea is that epistemic reasons are hypothetical reasons, and that this is enough to distinguish belief from other cognitive attitudes (this, recall, is one of the motivations for normativism). According to Olson, "moral error theory recognizes the wrongness of torture relative to UN declarations, but rejects categorical reasons to comply with such declarations" (2011: 88). Olson's way of dealing with CG arguments is the epistemic counterpart to this: epistemic reasons are (in Olson's terminology) *immanent*, rather than *transcendent*.

To speculate somewhat, Mackie is likely to accept a strategy of this sort. After all, most of his *Ethics* is devoted to normative ethical theorizing; just because (according to Mackie) there are no categorical norms does not mean we need to give up on normative ethics. Terry Horgan and Mark Timmons describe Mackie's position as follows:

> Mackie argued that all affirmative moral sentences are false because they involve (so he thought) metaphysical commitments to ontologically 'queer' properties, [but] he did not advocate eliminating the use of moral concepts and moral discourse; rather, he went on to propose a normative ethical system based on a certain conception of human flourishing (2000: 143-44).

Olson seems to suggest the epistemic error theorist be committed to something like this strategy: just because there are no categorical epistemic reasons does not mean that there are no reasons to believe in accordance with one's evidence (for example), only that the reasons in question are hypothetical, not categorical.

---

[36] Olson prefers the terms 'immanent' and 'transcendent' rather than 'hypothetical' and 'categorical', respectively, for reasons we need not get into here (cf. 2011: 79-80). Transcendent norms are just those norms that apply to agents categorically and have reason-giving force which transcends agents' desires, ends, activities, etc. Immanent norms are those norms which imply only non-categorical reasons, and have reason-giving force which does depend on agents' desires or ends.

But this response seems to miss the mark, at least when it comes to the CG argument under present consideration – the argument which allows us a doxastic analogue of the argument from queerness. The reason is that the error theorist is still committed to the existence of beliefs. In the moral case, the conclusion of the argument from queerness is that our moral *judgements* are systematically and uniformly false. It follows that there are no moral facts because these judgements are false: judgements of the form "You ought to do *Y*" are always false, so there are no moral facts. But doesn't the idea that our moral judgements are systematically and uniformly false imply, or presuppose, that there *are* moral beliefs that can be false? This presents no problem for the *moral* error theorist, since the facts about which moral error theorists are sceptical are moral facts: they are sceptical about the normative warranting relation between, for example, the suffering caused by torture and the wrongness of torture. But it seems the doxastic error theorist cannot concede that doxastic judgements express beliefs without undermining her own position. If the doxastic error theorist's claim were only that there are no categorical doxastic reasons, then there would presumably be no problem here. Like Olson, the doxastic error theorist could maintain that doxastic judgements only imply hypothetical reasons. But that is not the doxastic error theorist's claim, at least if we assume normativism. If we assume normativism, then part of the *concept* of belief is that belief is norm-governed; but a doxastic analogue of the argument from queerness would imply that nothing satisfies the norms which purport to be constitutive of belief, and so imply that there are *no* beliefs or believers. The doxastic error theorist, then, cannot coherently maintain that doxastic judgements are cognitive, expressing beliefs about beliefs, without undermining her own position. For example, according to the doxastic error theorist, the doxastic judgement "Sarah believes that snow is white" expresses the belief that Sarah believes that snow is white, i.e. it expresses a second-order belief. But our doxastic analogue of the argument from queerness concludes that doxastic judgements are systematically and uniformly false, which implies that there are *no* beliefs. It seems incoherent or self-

defeating for the doxastic error theorist to maintain that doxastic judgements express beliefs and at the same time endorse an argument which implies that there are no beliefs.

The foregoing objection to the doxastic error theory seems to be an instance of the argument from self-defeat, discussed above. Nishi Shah gives a similar argument against the doxastic error theorist, which takes the form of a dilemma:

> The error theorist about doxastic reasons faces a dilemma. He claims that all normative judgements attributing doxastic reasons are false, that there are no normative truths that would make some of our normative judgements about reasons for belief true. He must decide whether these normative judgements express beliefs or not. If he decides that they do not, then he must admit that normative statements about doxastic reasons, contrary to surface appearances, do not express propositions. If normative statements about reasons for belief do not express propositions, then it is not the case that they are false, contrary to the error theorist's central contention. But if the error theorist decides that normative judgements about doxastic reasons express beliefs, and attributing a belief entails making a normative judgement, he cannot consistently claim that there are no reasons for belief (2011: 102).

To make things simple, we can represent Shah's dilemma for the doxastic error theorist as follows:

(a) *Either* the doxastic error theorist holds that normative judgements express beliefs, in which case she cannot coherently claim that there are no *reasons* for belief, since she accepts that belief attribution involves making normative judgements;

(b) *Or* the doxastic error theorist denies that normative judgements express beliefs, in which case doxastic judgements are not false, since if normative judgements do not express beliefs or propositions, then they are not apt for assessment in terms of truth and falsity.

If the doxastic error theorist takes option (a), then she seems to undermine her own position – she succumbs to the argument from self-defeat because she maintains that normative judgements express beliefs (cognitivism), while accepting

that belief attribution involves making a normative judgement and denying that there are any (normative) reasons for belief. Her position becomes self-defeating.

If, on the other hand, the doxastic error theorist takes option (b), then she too faces difficulties. For one thing, she again seems to undermine her own position, because the conclusion of the doxastic argument from queerness is that doxastic judgements are systematically and uniformly false. But she has to deny that normative judgements express truth-apt propositions, which means doxastic judgements *cannot* be false (nor can they be true). But this brings us into a second and (perhaps) more significant issue: if the doxastic error theorist takes option (b), then she seems to have undermined what is distinctive about her brand of metaethical anti-realism – its *cognitivism*. Error theorists are cognitivists: they hold that the relevant judgements express propositions, but that they are all of them false. However, if the doxastic error theorist denies that normative judgements express propositions, then her position is no longer recognizable as

an error theoretic one. Rather, it would seem to be a form of doxastic non-cognitivism.[37]

If all this is right, then the moral error theory looks implausible: if the argument from queerness is sound, then – assuming normativism – a doxastic analogue of the argument from queerness can be developed. But this has an implausible upshot: the doxastic argument from queerness implies that there are no beliefs and no believers, and to believe that there are no beliefs and no believers would be to commit cognitive suicide. And this worry cannot be assuaged by appeal to the idea that doxastic normativity involves only hypothetical reasons, since even the doxastic error theorist needs to concede that there are beliefs and that normative judgements express beliefs, otherwise her position becomes either self-defeating or indistinguishable from non-cognitivism. I have, however, been assuming that normativism is true throughout this Chapter. It should be relatively clear that whether the foregoing is right depends largely on whether normativism is indeed true. But it remains to be seen that it is. As we shall see

---

[37] According to Frank Jackson (2000), the combination of non-cognitivism and the normativity of belief also yields the unpalatable consequence that there are no beliefs and no believers. The idea here is that, if non-cognitivism is true, then there is no such thing as *satisfying* normative constraints, where satisfying constraints is a matter of having the relevant properties. For example, being enrolled in university courses is a constraint on being awarded credit for those courses. To satisfy this constraint is to have the property of being such that you are enrolled in those courses. But there seem to be normative constraints on belief: if you believe that p, and that p implies q, then either you ought to believe that q, or you ought to revise your belief that p or that p implies q. To satisfy the relevant constraint in this situation is to have the normative property of being such that you ought to believe that q, or revise one of your antecedent beliefs. This, according to Jackson, poses a problem for non-cognitivism: "if non-cognitivism is true, there are no normative properties to have, or fail to have. Non-cognitivism is precisely the view that, although the language of normativity is meaningful, there are no properties corresponding to normative predicates. Equivalently, non-cognitivism about normativity is the view that normative predicates when attached to subject terms do not serve to make claims about how the subjects are; in consequence, there is no such thing as being how things are claimed to be by normative predicates and normative language in general. But then there is no such thing as subjects being how normative predicates say they are, and so no satisfying the normative constraints on being a believer. By Modus Tollens, therefore, we reach the conclusion that there are no believers and no beliefs" (2000: 102-3). This looks very much like a companions in guilt argument against non-cognitivism, though I cannot go into the details here. I take it that, if Jackson is right, then this makes option (b) in Shah's dilemma highly unattractive for the doxastic error theorist. And if she takes option (a), her position seems to become self-defeating. Either way, the doxastic error theorist is in trouble.

in the remaining two Chapters, normativism faces significant difficulties. I will argue that at least some of the main difficulties normativism faces can be met.

1.6 CONCLUSION

We have seen that, according to the moral error theorist, there are no moral facts: moral judgements are systematically and uniformly false. CG arguments (by analogy) typically purport to show either that the moral error theorist's scepticism about moral facts is unwarranted, or that arguments for the error theory are insufficient because they threaten to support an implausible error theoretic conclusion in other areas of discourse. I have argued that, assuming normativism, we can develop a doxastic analogue of the argument from queerness, and this has the paradoxical upshot that there are no beliefs and no believers. The argument from queerness, therefore, looks unsound, because if it were sound, it would support such an error theoretic conclusion about doxastic facts. This, of course, assumes that normativism is true, but it remains to be seen that it is.

2 TRUTH NORMS AND KNOWLEDGE NORMS

2.1 INTRODUCTION

In the last chapter, we saw that the threat to the moral error theorist is that, if normativism is true and the argument from queerness succeeds, then this implies an implausible doxastic analogue of the argument from queerness. This companions in guilt argument, however, relies on the assumption that normativism is true. In this chapter, I defend normativism against influential objections from Krister Bykvist and Anandi Hattiangadi. §2.2 clarifies what normativism amounts to, and what it means for a concept to be constitutively normative. Following Paul Boghossian's presentation of the theory, we shall see that a concept is constitutively normative just in case it is a condition of grasping the concept that one grasps statements explicable in normative or deontic vocabulary. I also explore some of the motivations for normativism. In §2.3, I present Bykvist and Hattiangadi's objections to the idea that belief is constitutively normative. According to them, truth norms for belief cannot be properly formulated without unpalatable consequences: truth norms are either false or impose unsatisfiable requirements on believers. In §2.4, I argue that *knowledge*, rather than merely *truth*, is the aim and norm of belief. If this is right, then three morals can be drawn from our story about the knowledge norm: (i) our version of the knowledge norm does not succumb to the objections Bykvist and Hattiangadi raise against truth norms; (ii) the knowledge norm is a plausible candidate for a constitutive doxastic norm; and (iii) unlike the truth norm, the knowledge norm does not clash with the less controversial epistemic norms in the 'hard cases' – lottery and coin-toss cases, in particular. The knowledge norm thus allows us to preserve the virtues of the truth norm, while avoiding its unpalatable consequences.

2.2 BOGHOSSIAN ON THE NORMATIVITY OF BELIEF

According to Paul Boghossian (2003, 2005), the concept of *mental content* is constitutively normative. A type of judgement is constitutively normative if and only if it is impossible that one understands judgements of that type without understanding statements that imply statements involving normative or deontic vocabulary ('ought', 'may', or 'reason'). So a concept is constitutively normative if grasping the statements in which it features requires one to grasp *oughts*, *mays*, or *reasons*. Boghossian's argument for the normativity of content has two parts. First, Boghossian argues that the concept of *belief* is constitutively normative. Normativism, we have seen, is the view that it is constitutive of, or essential to, belief that it is genuinely normative. By 'genuinely normative', I mean that the concept in question is *prescriptive* or *action-guiding*. Second, Boghossian argues that belief enjoys a certain conceptual primacy, so that the normativity of belief implies the normativity of mental content more generally.[38] Let us look only at the first part of the argument. To begin with, consider the following proposition concerning doxastic correctness:

(a) It is correct to believe that Mallory reached the summit if and only if Mallory reached the summit.

According to the normativist, a belief is correct if and only if the proposition believed is true.[39] We can express this in general terms with the following correctness principle:

---

[38] Since my focus here is primarily on the normativity of *belief*, rather than the normativity of content proper, I will not go into the details of Boghossian's argument for the conceptual primacy of belief. See Miller (2008) and Glüer and Wikforss (2009) for criticism of this strategy.

[39] More precisely, a belief is correct iff its *propositional content* is true. It should be relatively uncontroversial that beliefs are psychological states with contents that can be true or false; when I say that belief is truth-apt (i.e., apt for assessment in terms of truth and falsity) I mean that a belief is true when its content is true, and false when its content is false. It is not obvious that *truth* is a normative concept, and this is why many normativists about belief hold that 'correct' does not simply mean 'true'. As Nishi Shah puts it, for example, "beliefs are normatively assessable as correct or incorrect in virtue of the non-normative property that their propositional contents have of being true or false" (2011: 100).

(CORRECT): For any proposition p, the belief that p is correct if and only if p is true.

Correctness, according to Boghossian and other normativists,[40] is a normative notion: "[I]t seems right to say both that correctness is a normative matter, a matter of whether one ought to do what one is doing, and that the correctness conditions of one's thought are *constitutive* of what one is thinking" (Boghossian 2003: 35). Allan Gibbard makes a similar point:

> For belief, correctness is truth. Correct belief is true belief. My belief that snow is white is correct just in case the belief is true, just in case snow is white. Correctness, now, seems normative [...] The correct belief, if all this is right, seems to be the one [the subject] ought, in this sense, to have (Gibbard 2005: 338-9; cf. Gibbard 2012: Chapter 4).

The idea that correctness is a normative notion has intuitive pull: when we say that a belief is incorrect, we might mean that it is in some sense *defective*, or that it should be revised. Indeed, the fact that we can criticize others for having incorrect beliefs seems to require us to presuppose that, pre-theoretically, when someone believes a falsehood, they are doing something *wrong*. Now, if correctness is indeed a normative notion, then (a) seems to yield:

(b) One ought to believe that Mallory reached the summit if and only if Mallory did reach the summit.

Once again, we can express this in general terms as a norm of truth for belief:

(TRUTH): For any subject S and proposition p, S ought to believe that p if and only if p is true.[41]

Now, Boghossian's argument for the normativity of belief is contained partially in his response to Gibbard (2003). According to Gibbard, the difficulty with (b)

---

[40] See, e.g., Wedgwood (2002), Shah (2003), Shah and Velleman (2005), Engel (2007), Whiting (2010). For criticism of the idea that correctness is a normative notion, see Hattiangadi (2007) and Bykvist and Hattiangadi (2013).

[41] To keep matters simple, I will omit "For any subject S and proposition p" from the various iterations of doxastic norms in what follows. If such norms are in force, they should be thought of as *necessarily* so.

and norms like (TRUTH) is that the *ought* involved seems to express an *objective* norm, rather than a *subjective* one; and if we suppose that objective norms are genuinely constitutive of mental content and that they are genuinely normative, then this seems to yield the highly implausible result that *all* facts are normative. First, why is (b) an objective norm? Objective norms are those norms that hold in virtue of what is the case; it is not always *transparent* to the subject how she might follow them. Suppose I am thinking about the exact number of blades of grass per square foot in my garden, and suppose that there are, in fact, 4,871 of them. The belief that there are 4,871 blades of grass per square foot in my garden is the *correct* belief to have. But given (TRUTH), the belief that there are 4,871 blades of grass per square foot in my garden is the one I *ought* to have: I ought to have it only in virtue of what is, in fact, the case. But this is quite different from the claim that I ought to have this belief given the available evidence – indeed, I have no idea how many blades of grass are in my garden, and apart from counting each of them, I would have no way to know this. (TRUTH) and (b), then, are objective norms: if Mallory reached the summit, then I ought to believe that he did, regardless of whether I have any evidence that he did.

By contrast, subjective norms hold in virtue of considerations *available to the subject*; they are subjective only insofar as they engage with the subject's perspective. These considerations include facts about the evidence the subject has, the consistency of her beliefs, what it is rational for her to believe, and so on.[42] Consider, for example, the following subjective norms:

---

[42] To be clear: the objective/subjective distinction in this context has nothing whatsoever to do with the question of whether there are objective normative facts. An objective doxastic norm is simply one that holds in virtue of what is the case: I ought to believe that *p* because it is the case that *p*. Subjective epistemic norms are subjective only insofar as they engage with the subject's perspective: I ought to believe that *p* because I have overwhelming evidence that *p*, because it is rational for me to believe that *p*, and so on. I will say more below about cases in which the objective and the subjective norms can clash, e.g. cases in which the evidence misleads one to believe a falsehood.

(EVIDENCE): *S* ought to believe only that which is supported by the available evidence.

(CONSISTENCY): *S* ought to believe that (*p* and *q*) only if *p* and *q* are consistent.

(RATIONALITY): *S* ought to believe that *p* only if it is rational for *S* to believe that *p*.

And so on. Gibbard's worry is that if (b) is an objective norm, then if the 'ought to believe' is genuinely normative and constitutive of the belief that Mallory reached the summit, then this yields the unpalatable result that *all* facts are normative. For according to Gibbard, there is an *analytic equivalence* between both sides of the biconditional: the left-hand side purports to describe a normative fact and so, given analytic equivalence, the right-hand side *also* describes a normative fact. According to (b), therefore, "One ought to believe that Mallory reached the summit" and "Mallory did reach the summit" both describe normative states of affairs. The problem is that the second claim quite clearly describes a *non-normative* state of affairs, and so Gibbard rejects this line of reasoning.[43]

Now, Boghossian objects that Gibbard's argument rests on two questionable assumptions. First, Gibbard assumes that the relation between correctness conditions and normativity must take a biconditional form; and second, he assumes that constitutivity is a matter of analytic implication (Boghossian 2003: 36). Boghossian points out that (b) can be broken down into two conditionals:

(b*) One ought (objectively) to believe that Mallory reached the summit only if Mallory reached the summit.

---

[43] Note that Gibbard *is* a normativist, only he locates the normativity of content in the relation between content and *rules*, whereas Boghossian wants to locate it in the relation between content and *correctness*. For a sustained defense of his most recent view of the matter, see Gibbard (2012), though see, e.g., Wikforss (2018) for criticism.

(b**) If Mallory reached the summit, then one ought (objectively) to believe
     that Mallory reached the summit.

According to Boghossian, (b*) and (b**) are not on a par: the first claim, gener-
alized, says that one ought to believe only what is true, whereas the second gen-
eralization says that one ought to believe *everything* that is true. It is not the
case that for every fact about correctness, there is a corresponding *ought* fact,
for 'ought' implies 'can', while 'correct' does not have the same implication. More
precisely, (b**) violates the following intuitive and widely accepted principle:

('OUGHT' IMPLIES 'CAN'): If S ought to φ, then it is possible for S to φ.[44]

Boghossian writes: "although it is true that, for any p, if p, then it is correct that
p, it doesn't follow that if p, one ought to believe that p, for it's clearly impossible
to believe *everything* that's true […] I'm inclined to hold, therefore, that one can
infer p from 'One ought to believe that p,' but not the other way round" (2003:
37). This is Boghossian's objection to Gibbard's first assumption.

Gibbard's second worry is that a correctness-based normativity thesis will
overgeneralize: if true, it will prove that *all* facts are normative – it will prove
that the clearly non-normative claim "Mallory reached the summit" is, in fact,
normative. But Boghossian suggests this problem only arises given an improper
account of constitutivity, the account according to which a concept is constitu-
tively normative when it analytically implies normative statements, i.e. state-
ments involving deontic modalities such as obligation and permission. Regard-
ing *semantic* normativity, Gibbard writes:

> If meaning is "fraught with ought", the point is not just supposed to be the obvious one
> that *oughts* can depend on meanings. *Oughts* after all, can depend on all sorts of things
> that are clearly natural. Whether we ought to take a walk can depend on the weather;

---

[44] Although ('OUGHT' IMPLIES 'CAN') is widely accepted, it is not entirely uncontroversial. For
instance, Engel (2007) questions whether the principle is always true; and Mizrahi (2012) argues
that, from an epistemic point of view, ('OUGHT' IMPLIES 'CAN') is subject to counterexamples. I
cannot explore these issues here, so I shall assume in what follows that norms like (b**) are false
since they prescribe believing all the truths, including ungraspably complex ones.

that doesn't make the weather normative in any philosophically special sense. The claim of normativity is that *oughts* are somehow constitutive of meaning, that what something means *is a matter of* certain *oughts*' obtaining (2003: 84-5).

Gibbard's idea, according to Boghossian, is this: "q is constitutive of p just in case p analytically implies q – i.e., just in case facts about the meaning of 'p' and the meaning of 'q', along with logic, suffice to derive q from p" (Boghossian 2003: 33-4). However, Boghossian argues that this is not the proper account of constitutivity. Rather, constitutivity here should be understood as follows: q's being constitutive of p means not that p analytically implies q, but that it is a condition on *understanding* what it is for p to obtain that one understands what it is for q to obtain. To say that q is constitutive of p is to say that understanding or grasping the *concept* of a p-fact requires that one understand or grasp the *concept* of a q-fact. Note that this account is *stricter* than the analytic implication account:

> [T]he concepts ingredient in *Snow is white or grass is green* are sufficient for seeing that *Snow is white or grass is green* follows from *Snow is white*. But *Snow is white or grass is green* needn't be constitutive of *Snow is white* because it needn't be a condition on understanding *Snow is white* that one have the concept *or* or that one have the concept *grass* or, for all I know, that one have the concept *green* (Boghossian 2003: 38).

Mental content, according to Boghossian, is constitutively normative in this sense: content is constitutively normative if and only if it is a condition on understanding a particular content-attribution that one understands that, if the content attribution is true, then the relevant normative claim is true. The idea, for example, is that it is impossible that Marco understands what it is for Ebenezer to believe that Mallory reached the summit *without* also understanding that Ebenezer *ought* to believe that Mallory reached the summit only if Mallory did reach the summit.

Now, (b*) still seems like an objective norm, and it will not always be clear how one might follow it (more on this below and in Chapter 3). But the fact that this is not clear does not mean that (b*) does not express a genuine norm. Indeed, according to Boghossian "the holding of this norm is one of the defining features

of belief: it's what captures the idea that it is constitutive of belief to aim at truth. The truth is what you ought to believe, whether or not you know how to go about it, and whether or not you know if you have attained it" (2003: 38). Furthermore, Boghossian holds that the *subjective* epistemic norms are grounded in the objective ones. For example, it may be the case that we ought to believe what only what is supported by the evidence we have. The objective norm says that one ought to believe only what is true, and there may be subjective norms to believe only what one has evidence for. But, intuitively, beliefs for which there is good evidence are more likely to be true (as are beliefs that it is rational to hold, beliefs that are consistent, etc.), so by hypothesis, subjective norms such as norms of evidence will *typically* track what, according to the objective norm of truth, we ought to believe.[45]

Thus an important motivation for defending truth norms is that they *explain* why the less controversial subjective epistemic norms hold: the objective norm of truth explains why we ought to seek evidential justification for our beliefs, why our belief formation ought to be rational, why our beliefs ought to be consistent, and so on. Nearly all normativists in this debate endorse some version of this strategy.[46] Indeed, a common proposal is that the objective truth norm guides our belief formation and revision *indirectly*, via the subjective epistemic norms. As Boghossian puts it:

> [T]he "objective" norm that one ought to believe only what's true […] is not a rule that can be followed directly, but that can only be followed by following certain other rules, the so-called norms of rational belief. For example: that we ought to believe that which is supported by the evidence and not believe that which has no support; that we ought not to

---

[45] There will, however, be cases of conflict – cases in which the subjective norms prescribe believing what subjects (objectively) ought *not* to believe, i.e. falsehoods. For evidence can be misleading: plausibly, there are cases in which you have good evidence for the truth of a proposition, but that proposition turns out to be false. More on this in §2.4 below.

[46] In addition to Boghossian's papers (*op. cit.*), see Wedgwood (2002: 276; 2013a: 124; 2013b: 219 ff.), Shah (2003: 471), Millar (2004: 43 ff.), Shah and Velleman (2005: 519-20), Engel (2007: 181; 2013a: 36; 2013b: 207), Whiting (2010: 220-22; 2012: 289 ff.; 2013: 122), McHugh (2012b: 9), and Raleigh (2013: 256-7).

believe *p* if some alternative proposition incompatible with *p* has a higher degree of sup-
port; that we ought to believe *p* only if its degree of support is high enough, given the sort
of proposition that it is. And so on (2005: 101).

If normativism can be defended, therefore, we will have a neat picture of what
constitutes belief and why following norms of evidence, rationality, and so on is
something we ought to do. Moreover, doxastic norms – if they can be shown to
be constitutive of belief – will explain why false beliefs are faulty or defective:
they are faulty or defective because holding them violates the norm. Let us look
first at one influential argument *against* normativism.

2.3 TRUTH NORMS AND THEIR LOGICAL FORMS

Krister Bykvist and Anandi Hattiangadi ('B&H' hereafter) object to arguments
of the sort made by Boghossian on the grounds that truth-norms cannot be
properly formulated without unpalatable consequences (2007, 2013). Let us now
examine this line of argument, which we might call *the formulation problem*.
Recall that Boghossian is committed to the claim that a belief's correctness lies
in its truth-conditions. If correctness is an essentially normative notion, then
(TRUTH) follows from (CORRECT). Boghossian showed that the norm I am calling
(TRUTH) breaks down into two conditional claims and argued that (b*) is the
more plausible one – the one that is constitutive of belief. B&H want to reject all
such formulations on the grounds that their requirements are unsatisfiable.
They argue, first, that (TRUTH) is *ambiguous* in that the *ought* which features in
it can take either a *wide* or a *narrow* scope over the biconditional. The narrow
scope reading (TRUTH$_N$) and the wide scope reading (TRUTH$_W$) can be formulated
as follows:

(TRUTH$_N$): *S* ought to (believe that *p*) if and only if *p* is true.

(TRUTH$_W$): *S* ought to (believe that *p* if and only if *p* is true).

The narrow scope reading requires us to believe that *p* on the condition that *p* is
true. The wide scope reading, by contrast, requires us unconditionally to respect

the entire biconditional. Let us first examine the narrow scope reading. As with (b) above, (TRUTH$_N$) can be broken down into two conditionals:

(TRUTH$_N$*): If p is true, then S ought to (believe that p).

(TRUTH$_N$**): If S ought to (believe that p), then p is true.[47]

Now, given ('OUGHT' IMPLIES 'CAN'), (TRUTH$_N$*) is highly implausible, since it says that for *any* true proposition, one ought to believe it. But there are presumably infinitely many propositions – probably most of them trivial – and there are propositions which are simply too complex for a human being to believe. Suppose that p is a proposition far too complex for any human being to believe. Now, insofar as 'ought' implies 'can', it cannot be the case that S ought to believe that p, so (TRUTH$_N$*) looks highly implausible, if not manifestly false.

What about (TRUTH$_N$**)? This, recall, was the norm that Boghossian ends up endorsing. B&H object that it is not obvious that (TRUTH$_N$**) captures the intuitions Boghossian thinks it captures: whereas (TRUTH$_N$*) "places too demanding a requirement on believers, [TRUTH$_N$**] places no requirement at all" (2007: 280). They explain:

> Obviously, if p is true, nothing whatsoever follows from [TRUTH$_N$**] about what S ought to believe. Less obviously perhaps, if p is false, nothing whatsoever follows about what S ought to believe. For, if p is false, it only follows that it is not the case that S ought to believe that p. It does not follow, from the falsity of p, that S ought not to believe that p. There is an important difference between 'it is not the case that S ought to believe that p' and 'S ought not to believe that p' – the former states that S *lacks* an obligation to believe that p and the latter states that S *has* an obligation not to believe that p. The former is compatible with its being permissible for S to believe that p, while the latter is incompatible with its being permissible for S to believe that p. Hence, whether p is true or false, [TRUTH$_N$**] does not tell S what to believe (2007: 280).

[47] Alternatively: S ought to (believe that p) only if p is true.

(TRUTH$_N$) is simply the conjunction of (TRUTH$_N$*) and (TRUTH$_N$**). Since, according to B&H, (TRUTH$_N$*) is false and (TRUTH$_N$**) is too weak, (TRUTH$_N$) itself is untenable.

Furthermore, B&H dispute a proposal from Ralph Wedgwood[48] which seems, initially, to dissolve the objection according to which the truth norm violates ('OUGHT' IMPLIES 'CAN'). Wedgwood suggests that we restrict the true propositions one ought to believe to the propositions which one has *actually considered*:

(TRUTH*): If S considers whether p, then S ought to (believe that p) if and only if p is true.

The problem with (TRUTH$_N$*) was that it entails that we ought to believe all the truths (of which there are infinitely many), including truths which are too complex for us to believe. Insofar as 'ought' implies 'can', it cannot be the case that we ought to believe everything that is true. But (TRUTH*) is more plausible: intuitively, it is possible to believe propositions which one has actually considered; but it is not possible to consider infinitely many propositions, nor is it possible to consider propositions too complex for human beings to believe. That one can believe only what one can consider seems, *prima facie*, to be a psychological fact about human beings. If you consider a true proposition, then you ought to believe it; if you consider a false proposition, you ought not to believe it.

B&H dispute this suggestion as follows. There are some propositions – Moore-paradoxical propositions, or 'blindspots' as Sorensen (1988) calls them – which it is logically impossible to believe truly: "if they are true, then you don't believe them, and if you believe them, then they are false" (B&H 2007: 281). Standard examples of blindspot propositions include the following:

It is raining and nobody believes that it is raining.

---

[48] Wedgwood (2002) and personal communication with Bykvist and Hattiangadi (2007: 280). See Wedgwood (2013a) for a reply.

There are no believers.

If it is true that there are no believers, then you do not believe that there are no believers; if it is true that it is raining but nobody believes that it is raining, it follows that you do not believe that it is raining. Likewise, if you believe these propositions, then they are false. Now, Wedgwood's suggestion will not help here, for according to B&H, one can consider whether it is raining and nobody believes it is raining, or whether there are no believers, but it is not the case that these propositions can be *truly* believed. B&H note that the problem is not that the propositions in question cannot be believed, but rather that the obligation to believe p iff p is true cannot be satisfied. The objection is not that 'ought' implies 'can', but that 'ought' implies 'can satisfy' (B&H 2007: 282):

('OUGHT' IMPLIES 'CAN SATISFY'): If S ought to φ, then it is possible for S to φ while still being obligated to φ.[49]

Why does (TRUTH*) violate ('OUGHT' IMPLIES 'CAN SATISFY')? Suppose that p is the proposition "It is raining and nobody believes that it is raining". Now, if you believe that p, then p is false, in which case you are no longer obligated to believe that p. So, if one believes true blindspots, it is not possible to φ while *still* being obligated to φ. In other words, in cases in which you consider a proposition, you ought to believe it iff it is true. But it is impossible that you believe a true blind-spot while still being obligated to believe true blindspots. For if you believe blind-spots, then they are false, and so *it is not the case* that you ought to believe them.

Now, B&H consider and reject an alternative way to formulate Wedgwood's suggestion which purports to accommodate blindspots. Consider the following modification to (TRUTH*):

(TRUTH**): If S considers whether p, and p is truly believable, then S ought to (believe that p) if and only if p is true.

---

[49] See Bykvist (2007) for discussion of this principle in an ethical context.

If p is the proposition "It is raining and nobody believes that it is raining" and it is true, then (TRUTH**) does not imply that S ought to believe that p, since p is not truly believable. However, B&H claim that (TRUTH**) is too weak: it does not tell S that she ought *not* to believe that it is raining and nobody believes that it is raining. For even if p is true – as we have supposed – (TRUTH**) tells us nothing about what we ought *not* to believe. But the intuitive response to p would be that S ought not to believe that p, even if it is true. Furthermore, B&H claim that (TRUTH**) conflicts with a similarly plausible principle:

(TRUTH***): If S considers whether p, and p is not truly believable, then S ought not to (believe that p).

It is possible, according to B&H, that there are true propositions which are not truly believable but which consist of *conjuncts* which are themselves truly believable. Suppose that S considers the proposition "It is raining and nobody believes that it is raining" and that this proposition is true. According to (TRUTH***), S ought not to believe that it is raining and nobody believes that it is raining, since this proposition is not truly believable. But considering a conjunction requires that one consider each of its conjuncts: if the conjunction is true, each of its conjuncts must be true, and so even if the conjunction is not truly believable, each of the conjuncts *is* truly believable (B&H 2007: 282-3). Now, according to (TRUTH**), S ought to believe the proposition "It is raining" and the proposition "Nobody believes that it is raining". But if we combine (TRUTH**) and (TRUTH***), it follows that S ought to believe the conjunct "It is raining" and the conjunct "Nobody believes that it is raining", but S ought *not* to believe the conjunction "It is raining and nobody believes that it is raining". More formally, the combination of (TRUTH**) and (TRUTH***) violates the following intuitive principle:

(AGGLOMERATION): If S ought to (believe that p) and S ought to (believe that q), then S ought to (believe that p and q).

B&H note that even though (AGGLOMERATION) is intuitively plausible, it is not entirely uncontroversial. "However," they write, "many of those who defend the normativity of belief have reason to accept this principle, since they also defend the normativity of content […] If content is constituted by norms, it is plausible that [AGGLOMERATION] will be constitutive of the ordinary concept of *and*" (2007: 283). The idea is that (AGGLOMERATION) is a plausible candidate for a 'normativized' version of the conjunction-introduction rule in propositional logic (B&H 2013: 111).[50]

Moreover, even if the normativist denies (AGGLOMERATION), she still needs to accept the existence of *doxastic dilemmas* – scenarios in which subjects believe the propositions that they ought to believe but end up in situations in which they believe the propositions they ought not to believe in that situation. Suppose that it is raining, and that S considers the proposition that it is raining. Suppose again that the proposition that it is raining is a proposition no one believes. Since it is true that it is raining, and since – we can assume – (TRUTH**) is also true, it follows that you ought to believe that it is raining. That is, from the truth of the following proposition (which S has considered):

It is raining.

Combined with (TRUTH**):

If S considers whether p, and p is truly believable, then S ought to (believe that p) if and only if p is true.

It follows that:

S ought to believe that it is raining.

Now, given (TRUTH**) and our supposition that no one believes that it is raining, it likewise follows that you ought to believe that no one believes it is raining. But: "if you satisfy the first obligation – to believe that it is raining – then, given

---

[50] Conjunction-introduction: p, q, therefore (p & q).

[TRUTH***] and the fact that you now believe that it is raining, you ought not to believe that no one believes that it is raining" (B&H 2007: 283). That is, from the truth of the following proposition (which S has considered):

No one believes that it is raining.

Combined with (TRUTH***):

If S considers whether p, and p is not truly believable, then S ought not to (believe that p).

It follows (since "No one believes that it is raining" is not truly believable) that:

S ought not to believe that no one believes that it is raining.

The first obligation (to believe that it is raining) and the second obligation (to believe that no one believes that it is raining) are satisfiable *separately*, but they are not satisfiable *jointly*: whereas each obligation on its own conforms to ('OUGHT' IMPLIES 'CAN SATISFY'), these obligations do not jointly conform to this principle. So, the normativist needs to accept the existence of this sort of doxastic dilemma: if you believe what you ought to believe, you end up in a situation in which you believe what you ought not to believe in that situation.

According to B&H, therefore, the narrow scope reading (TRUTH_N) is untenable. What of (TRUTH_W)? B&H note that there are two combinations of states of affairs which can satisfy the requirements of (TRUTH_W): either S believes that p and p is true, or it is not the case that S believes that p and p is false. Furthermore, there are two more combinations that should be avoided: either S believes that p and p is false, or it is not the case that S believes that p and p is true. Now, the same objections B&H lodge against (TRUTH_N) do not apply to (TRUTH_W) in any obvious way, since the *ought* in (TRUTH_W) takes wide scope over the biconditional and so cannot be broken down into two separate conditionals, such as (TRUTH_N*) and (TRUTH_N**).

B&H object to (TRUTH$_W$) on the grounds that it does not capture the intuition that the truth is what you ought to believe. We are unconditionally required by (TRUTH$_W$) to respect the entire biconditional, and so the only way for a believer to satisfy the norm is to satisfy one of the combinations noted above. But – since the *ought* takes wide scope – this means that we cannot *detach* from (TRUTH$_W$) the obligation to believe that p, even when p is true (cf. Broome 2000). By analogy, consider the following syllogism:

(i) S ought (if S believes that p and that p implies q, to believe that q)

(ii) S believes that p and that p implies q.

(iii) Therefore, S ought to believe that q.

Given the wide scope of the *ought* operator in (i), (iii) does not follow from (i) and (ii). That is, we cannot detach S's obligation to believe that q from (i) and (ii). B&H write:

> […] q might well be an absurd proposition, in which case what you ought to do is not believe that p, or not believe that p implies q. Similarly, it does not follow from [TRUTH$_W$] that you ought to believe that p, even when p is true, and [TRUTH$_W$] does not therefore capture the thought that the truth is what you ought to believe. Nor, for that matter, does [TRUTH$_W$] capture the thought that a false belief is defective. From the falsity of p and [TRUTH$_W$] it does not follow that you ought not to believe that p. [TRUTH$_W$] says that when you believe a falsehood, all is not as it ought to be, but this does not imply that it is the belief which is faulty or defective (2007: 284).[51]

Moreover, B&H argue that there are instances in which (TRUTH$_W$) does violate ('OUGHT' IMPLIES 'CAN') after all, "for values of p such that the truth of p is unavoidable, yet p is unbelievable" (2007: 284). Suppose that p is the conjunction of all necessary truths. Obviously, this would be far too complex for a human being to believe; but according to (TRUTH$_W$), you ought either to bring it about that p is false or bring it about that you believe that p. But this does not seem possible. Intuitively, the conjunction of all necessary truths is itself a necessary truth,

---

[51] See also Kalantari and Luntley (2013) for relevant discussion.

and so you cannot bring it about that p is false. Moreover, the conjunction of all necessary truths is presumably too complex for a human being to grasp, so you cannot believe that p either. Insofar as 'ought' implies 'can', then, B&H maintain that (TRUTH_W) is false.

2.4 KNOWLEDGE AS THE AIM AND NORM OF BELIEF

We have seen that B&H's formulation problem constitutes a significant challenge to the normativist. In this section, I want to suggest we hold that *knowledge*, rather than merely truth, is the aim and norm of belief.[52] This, I shall argue, might allow us to solve the formulation problem and preserve all the virtues of the truth norm, while avoiding its unpalatable consequences.

According to Hattiangadi (2010), there are two issues arising from the idea that belief has an aim: one epistemological, and one metaphysical. The epistemological issue is about the epistemic evaluation of belief: it is about whether truth, knowledge, or whatever is of fundamental epistemic value. I am primarily concerned, however, with the metaphysical issue, which is about the essential nature of belief. As we have seen, normativists interpret the slogan "Belief aims at X" as a *normative*, metaphorical claim. If, as per the present suggestion, knowledge is the aim of belief – and assuming we can give this claim a normative interpretation – then *what* is the knowledge norm for belief? Suppose we have the following knowledge norm:

(KNOWLEDGE): *S* ought to believe that p if and only if S knows that p.[53]

---

[52] Of course, I have not considered *all* the ways we might formulate the truth norm. For example, some philosophers have opted for a weaker deontic modality that 'ought' in formulating the truth norm (Whiting 2010, 2012, 2013). On this view, the truth norm is not about what we *ought* to believe, but rather what we are *permitted* to believe. For criticism of permissive truth norms, see McHugh (2012b), Bykvist and Hattiangadi (2013), and Greenberg (201X).

[53] Different versions of the knowledge norm are defended by Williamson (2000: Chapter 11), Sutton (2007), Smithies (2012), Gibbons (2013: Chapter 8), Engel (2013a), and Littlejohn (2013). According to Littlejohn (2013: 294), the difficulties Bykvist and Hattiangadi raise for truth norms do not seem to arise for knowledge norms. In fact, however, we shall see presently that many of these difficulties *do* arise for knowledge norms (but not, I think, all of them).

One might worry, however, that there is something strange about this norm, since knowledge entails belief: if you know that p, then you believe that p. There is something dubious about the left-to-right reading of the biconditional. However, let us postpone discussion of this. Perhaps a more pressing issue is that, like (TRUTH), (KNOWLEDGE) is ambiguous, since the *ought* operator can take either a narrow scope or a wide scope over the biconditional. The narrow scope reading (KNOWLEDGE$_N$) and the wide scope reading (KNOWLEDGE$_W$) run as follows:

(KNOWLEDGE$_N$): S ought to (believe that p) if and only if S knows that p.

(KNOWLEDGE$_W$): S ought to (believe that p if and only if S knows that p).

Just like (TRUTH$_N$), (KNOWLEDGE$_N$) prescribes believing that p on the condition that one knows that p. (KNOWLEDGE$_W$), on the other hand, obligates subjects to respect the entire biconditional. (KNOWLEDGE$_N$) is the conjunction of two conditionals:

(KNOWLEDGE$_N$*): If S knows that p, then S ought to (believe that p).

(KNOWLEDGE$_N$**): If S ought to (believe that p), then S knows that p.[54]

On the face of it, (KNOWLEDGE$_N$*) is attractive, since unlike (TRUTH$_N$*), it does not prescribe believing everything that is true, nor does it prescribe believing ungraspably complex truths. This is so precisely because we cannot *know* all the truths, and we cannot have knowledge of ungraspably complex ones. (Granted, (KNOWLEDGE$_N$*) will prescribe believing *some* trivial truths, because we can have knowledge of trivial truths.) Thus (KNOWLEDGE$_N$*) does not violate ('OUGHT' IMPLIES 'CAN').

However, (KNOWLEDGE$_N$*) faces difficulties. For one thing, it is not possible to *violate* the norm, since knowledge entails belief: if you know that p, then you

---

[54] Alternatively: "S ought to (believe that p) only if S knows that p".

believe that p. But any plausible norm implies the possibility of violation: normative *force* goes hand-in-hand with normative *freedom* (Railton 2000). In other words, just as 'ought' implies 'can', 'ought' implies 'can not':

('OUGHT' IMPLIES 'CAN NOT'): If S ought to φ, then it is possible for S to fail to φ.

Thus (KNOWLEDGE_N*) looks implausible.[55] What of (KNOWLEDGE_N**)? The problem with (TRUTH_N**) was that the norm is too weak, and so does not capture the intuition that the truth is what one ought to believe (B&H 2007: 280). A similar problem arises in the present case. Suppose S knows that p. Now, (KNOWLEDGE_N**) does not tell S whether she ought to believe that p. Consider the following (invalid) syllogism:

(1) If S ought to (believe that p), then S knows that p.
(2) S knows that p.
(3) Therefore, S ought to (believe that p).

Obviously, this argument affirms the consequent. So if it is the case that S knows that p, nothing follows from (KNOWLEDGE_N**) about what S ought to believe. Moreover, if S does *not* know that p, nothing follows about what she ought to believe either. For if it is false that S knows that p, all that follows is that *it is not the case* that S ought to believe that p. To make this clear, consider the following (valid) syllogism:

(1) If S ought to (believe that p), then S knows that p.
(2) It is not the case that S knows that p.
(3) Therefore, it is not the case that S ought to (believe that p).

---

[55] The norm also faces another difficulty. If S *doesn't* know that p, it doesn't follow from (KNOWLEDGE_N*) that S ought *not* to believe that p. The intuition behind the knowledge norm, however, is meant to be that beliefs which do not amount to knowledge are faulty or defective. Intuitively, this cannot be captured by a norm which doesn't tell you that you ought not to believe a proposition that you don't know.

As B&H point out in the case of the truth norm, there is a difference between "It is not the case that S ought to believe that p" and "S ought not to believe that p": according to the former, S *lacks* an obligation to believe that p, whereas according to the latter, S *has* an obligation *not* to believe that p. Thus (3) is compatible with it being permissible for S to believe that p. Like (TRUTH$_N$**), therefore, (KNOWLEDGE$_N$**) does not tell the subject what she ought to believe, regardless of whether she knows that p.

It seems to me, then, that both (KNOWLEDGE$_N$*) and (KNOWLEDGE$_N$**) are implausible. If that is right, then (KNOWLEDGE$_N$) is untenable, since this norm is simply the conjunction of (KNOWLEDGE$_N$*) and (KNOWLEDGE$_N$**). But is this fatal for the narrow scope reading of (KNOWLEDGE)? A natural next step for the normativist might be to modify the knowledge norm with the proviso that subjects *must have considered* the propositions they ought to believe:

(KNOWLEDGE*): If S considers whether p, then S ought to (believe that p) if and only if S knows that p.

On the face of it, (KNOWLEDGE*) is attractive, though it ultimately faces difficulties. First, here is why it is attractive. B&H, we have seen, lodge a 'blindspot' objection against (TRUTH*) – the alethic equivalent to (KNOWLEDGE*). Blindspot propositions are such that it is logically impossible to believe them truly: if they are true, then you don't believe them; if you believe them, then they are false. In the case of the truth norm, (TRUTH*) violates ('OUGHT' IMPLIES 'CAN SATISFY'): you cannot believe a true blindspot p and *still* be obligated to believe that p because, once you believe that p, p is false, in which case you lack the obligation to believe that p. But (KNOWLEDGE*) faces no such difficulty. It will never prescribe believing blindspots because *we cannot have knowledge of blindspots*. Suppose that p is the blindspot proposition "It is raining and nobody believes that it is raining". If you believe that p, then p is false, and so you do not know that p. Knowledge, after all, is factive (Williamson 2000); it entails truth. Moreover, if p is true, then you do not believe that p, and so it follows that you do not know

70

that p. For knowledge entails belief. (KNOWLEDGE*), therefore, will never prescribe believing blindspots because its consequent will never be satisfied in the case of blindspot propositions: if you consider some blindspot proposition p, then it is not the case that you know that p, so you will lack the obligation to believe that p.

But this is precisely the difficulty with (KNOWLEDGE*): when you do not know that p, all that follows from the norm is that you *lack* an obligation to believe that p. It does not follow that you *have* an obligation *not* to believe that p. Thus by modifying the knowledge norm with the proviso that subjects consider the proposition in question, we simply push the problem raised for (KNOWLEDGE$_N$**) back a step. Suppose that p is the proposition "It is raining and nobody believes that it is raining" and that S considers whether p. Now, we can illustrate the difficulty here with the following argument:

(1) If S considers whether p, then S ought to (believe that p) if and only if S knows that p.
(2) S considers whether p.
(3) Therefore, S ought to (believe that p) if and only if S knows that p.
(4) It is not the case that S knows that p.[56]
(5) Therefore, it is not the case that S ought to (believe that p).

But now it should be clear that (KNOWLEDGE*) faces essentially the same difficulty as (TRUTH$_N$**) and (KNOWLEDGE$_N$**), only the problem has been pushed back a step: when p is a blindspot, nothing follows from (KNOWLEDGE*) about what subjects ought to believe, even after they have considered the propositions in question.

So it seems as if the narrow scope reading of the knowledge norm is hopeless. What of the wide scope reading (KNOWLEDGE$_W$)? This, too, faces difficulties. Recall that in the case of (TRUTH), there are two combinations of states of affairs

---

[56] Because p is a blindspot proposition.

which satisfy the requirements of the norm: either you believe that p and p is true, or it is not the case that you believe that p and p is false. Moreover, there are two combinations of states of affairs which ought to be avoided: either you believe that p and p is false, or it is not the case that you believe that p and p is true. B&H, we saw, argued that (TRUTH$_W$) does not capture the intuition that the truth is what one ought to believe, nor the thought that false beliefs are faulty or defective. Since (TRUTH$_W$) deals in combinations, and since the *ought* takes wide scope over the biconditional, we cannot detach from (TRUTH$_W$) the obligation to believe that p, even when p is true.

A similar problem arises for (KNOWLEDGE$_W$). The wide scope knowledge norm requires subjects unconditionally to respect the entire biconditional. Thus there are two combinations of states of affairs which satisfy the norm: either S knows that p and she believes that p, or S does not know that p and it is not the case that she believes that p. And there are two combinations of states of affairs which ought to be avoided: either S does not know that p and she believes that p, or S knows that p and it is not the case that she believes that p.[57] But precisely because the norm deals in combinations, more or less the same difficulty arises for (KNOWLEDGE$_W$) as for (TRUTH$_W$): when you believe a proposition that you do not know, all is not as it ought to be, but we do not know whether it is the *belief* that is defective. Suppose, for example, that you believe that p without knowing that p. There are two ways to remedy this situation: either you give up the belief that p, or you acquire knowledge that p. The norm, however, does not enjoin one option rather than the other: we cannot be sure whether it is the *belief* which is defective, as opposed to one's lack of knowledge that p. Thus (KNOWLEDGE$_W$) does not capture the intuition that we ought to believe what we know, just as

---

[57] The latter combination ("S knows that p and it is not the case that she believes that p") is of course not an option; you avoid it by default. For knowledge entails belief: if you know that p, then you believe that p. So you never know that p without believing that p.

(TRUTH~W~) does not capture the intuition that the truth is what you ought to be-
lieve. Likewise, (KNOWLEDGE~W~) does not capture the thought that beliefs which
do not amount to knowledge are faulty or defective.

So (KNOWLEDGE~W~) looks implausible. Is there a way out for the normativist?
At the beginning of this section, I raised one potential worry for the knowledge
norm: knowledge entails belief, so if you know that p, then you *already* believe
that p. There is perhaps something odd about a norm which prescribes forming
a belief which the subject has already formed. The 'way out' for the normativist
that I want to suggest involves dealing with this issue. Perhaps it makes more
sense to talk about what we are *in a position to know*, rather than what we know,
in the context of doxastic norms. This, at least, will not entail that one has al-
ready formed the very belief the norm tells them they ought to form. Hattiangadi
(2010: 431) considers and rejects the following knowledge norm for belief:

> (KNOWLEDGE**): If S considers whether p, then
>
> (i) S ought to (believe that p) if she is in a position to know that p; and
>
> (ii) S ought not to (believe that p) if she is not in a position to know that p.

To say that you are in a position to know that p is to say that, if you were to
believe that p, you would know that p. But Hattiangadi lodges the following ob-
jection to this norm:

> The semantics of this conditional can be given in terms of possible worlds. We do not want
> to say that you are in a position to know that p just in case all possible worlds at which
> you believe that p are worlds at which you know that p. For, this condition would rarely
> be met, and you would rarely be in a position to know that p. Instead, it makes more sense
> to say that you are in a position to know that p just in case all of the closest worlds at
> which you believe that p are worlds at which you know that p. This means that even if you
> are in a position to know that p, it is still possible that you will believe that p, and yet fail
> to come to know that p. In such a situation, [KNOWLEDGE**] implies that you ought to
> believe that p. However, if when you believe that p you realize a world at which you do not
> know that p, then you will be in a state that is not epistemically ideal, although you will
> have formed your doxastic attitudes in accordance with [KNOWLEDGE**] (2010: 31).

But it is not clear that this objection succeeds. For it seems that, if when you believe that p you realize a world at which you do not know that p, you are in a state that is not epistemically ideal, but you have *violated* (KNOWLEDGE**): you have not, contra Hattiangadi, formed your doxastic attitudes in accordance with the norm. Suppose that Jones is in a position to know that p. That is, in all close possible worlds in which Jones believes that p, he knows that p. Now, (KNOWLEDGE**) implies that in some close possible world W_1, Jones knows that p. But, we can suppose, there is a *distant* possible world W_2 in which Jones believes that p but does not know that p. In W_2, Jones is in a state that is not epistemically ideal. Jones is not in a position to know that p because W_2 is a distant world (from the actual world and W_1), and in the closest possible worlds to W_2, Jones will not know that p. But it is not clear, then, that in W_2 Jones has formed his doxastic attitudes in accordance with (KNOWLEDGE**): since Jones is not in a position to know that p, the norm implies that Jones ought *not* to believe that p in W_2. So, contra Hattiangadi, we have a *violation* of (KNOWLEDGE**) in this case.

According to Hattiangadi, the problem with (KNOWLEDGE**) is that it expresses an obligation we have under certain conditions, rather than a *conditional obligation*. The idea is that knowledge norms are meant to capture the thought that knowing that p while believing that p is an epistemically ideal state, whereas believing that p without knowing that p is an epistemically defective state. Hattiangadi (2010: 431) argues that we can capture this with the following *wide scope* knowledge norm:

(KNOWLEDGE***): If S considers whether p, then S ought to (believe that p only if she knows that p).

(KNOWLEDGE***) tells us which states are *epistemically permitted* (believing that p and knowing that p) and *epistemically forbidden* (believing that p and not knowing that p). Hattiangadi suggests we be committed to (KNOWLEDGE***) as a *criterion of rightness*, rather than as a decision-making procedure (more on

this in Chapter 3). If we think of the norm this way, (KNOWLEDGE***) does not succumb to a 'detachment' objection, since it does not itself prescribe anything. Rather, the subjective epistemic norms do the prescribing; the objective criterion of rightness (KNOWLEDGE***) *sets the standard* for which epistemic norms to follow. However, if what I have said about (KNOWLEDGE**) above is correct, then we have been given no reason to prefer (KNOWLEDGE***) over (KNOWLEDGE**).

I take it that at least three morals can be drawn from the foregoing story. First, (KNOWLEDGE**) does not succumb to the objections Bykvist and Hattiangadi raise against truth norms, nor does Hattiangadi's (2010) objection to it succeed. Second, (KNOWLEGDGE**) is a plausible candidate for a constitutive doxastic norm that serves to distinguish belief from other cognitive attitudes. Finally, (KNOWLEDGE**) can explain why the less controversial epistemic norms hold: it explains, for instance, why we ought to seek evidential justification for our beliefs, why our belief formation ought to be rational, and so on. In short, it seems to me that (KNOWLEDGE**) preserves all the virtues of the truth norm while avoiding its unpalatable consequences. Let me defend each of these three claims.

To begin with, (KNOWLEDGE**) does not succumb to the objections B&H raise against truth norms. For one thing, (KNOWLEDGE**) does not violate ('OUGHT' IMPLES 'CAN'): it does not prescribe believing everything that is true, nor ungraspably complex propositions. There are two reasons for this. First, the norm is modified with the proviso that subjects *consider* the propositions in question. Since we cannot consider every proposition, the norm does not violate ('OUGHT' IMPLIES 'CAN'). Moreover, we are never in a position to know ungraspably complex truths, precisely because they are ungraspable. Thus the norm does not prescribe believing ungraspably complex truths. Nor does (KNOWLEDGE**) violate ('OUGHT' IMPLIES 'CAN SATISFY'). As argued above, we can never have knowledge of blindspot propositions. I take it that this entails that we are never *in a position to know* blindspots. For to say that S is in a position to know that p

75

is to say that, were S to believe that p, she would know that p. But if one believes a true blindspot p, then p is false, and so one would not know that p (knowledge entails truth). Thus (KNOWLEDGE**) does not prescribe believing blindspots, and so does not violate ('OUGHT' IMPLIES 'CAN SATISFY'). (Moreover, (KNOWLEDGE**) has a narrow scope *ought*, so it does not succumb to a 'detachment' objection.) The norm, therefore, does not succumb to the objections B&H raise against truth norms.

As for our second moral, (KNOWLEDGE**) seems a plausible candidate for a constitutive doxastic norm – one that serves to distinguish belief from the other cognitive attitudes. On the foregoing account of constitutivity, a concept is constitutively normative just in case it is a condition on grasping the statements in which it features that one grasps normative or deontic statements. It seems at least *prima facie* plausible that, in order to have:

> Ebeneezer understands that Marco believes that Mallory reached the summit.

We need something like the following:

> Ebeneezer understands that (if Marco considers whether Mallory reached the summit, then if Marco is in a position to know that Mallory reached the summit, then he ought to believe that Mallory reached the summit; and if Marco is not in a position to know that Mallory reached the summit, then he ought not to believe that Mallory reached the summit).

Though of course more needs to be said. However, it at least seems plausible that (KNOWLEDGE**) can distinguish belief from the other cognitive attitudes. For the following norms seem patently false (assuming S considers whether p, and simplifying somewhat):

> (IMAGINE): S ought to (imagine that p) iff S is in a position to know that p.

> (WISH): S ought to (wish that p) iff S is in a position to know that p.

(HOPE): S ought to (hope that p) iff S is in a position to know that p.

And so on. We can imagine, hope, or wish whatever we want: we can imagine that a state of affairs obtains (say) without asking ourselves whether or not we are in a position to know that the state of affairs does in fact obtain. Whether or not we are in a position to know that p has no bearing on whether or not we are permitted to imagine, wish, or hope that p. We might put the point in terms of *commitment* (cf. Gibbons 2013: Chapter 8). To believe that p entails commitment to p's being true, whereas when you imagine that p, you imagine that p is true, but you are not *committed* to p's being true. Likewise for wishing and hoping. Since knowledge entails both truth and belief, (KNOWLEDGE**) distinguishes belief from the other cognitive attitudes in this way: if you believe that p when you are in a position to know that p – and thereby acquire knowledge that p – you are committed to p's being true. For your knowledge that p entails commitment to p's being true *via* your belief that p.

Finally, (KNOWLEDGE**) can explain why the subjective epistemic norms hold. It is worth spending more time on this. As noted previously, a number of philosophers have pointed out that the strategy according to which norms like (EVIDENCE) are *explained* by the objective norm of truth is problematic, since the objective norm can clash with the subjective ones in some cases.[58] Evidence, for example, can be misleading: you might have excellent evidence for a falsehood. Norms like (EVIDENCE), therefore, cannot be explained in terms of the truth norm in all cases, and this casts doubt on the idea that the objective norm of truth explains why the subjective epistemic norms hold – or to use Boghossian's terminology, that it supplies their 'rationale'. Of course, following evidential norms will *generally* lead to the truth; it is just that they cannot lead to the truth in all cases, because there are some instances in which following the epistemic norms will violate the truth norm.

---

[58] In this section, I will focus on Hattiangadi's (2010) presentation of the problem. I discuss the objections of Glüer and Wikforss (2013) in the next chapter. See also Gibbons (2013).

First, let us examine in more detail the idea that the objective norm of truth can clash with the epistemic norms, before looking at the way in which (KNOWLEDGE**) helps to deal with this problem. To begin with, consider the alethic counterpart to (KNOWLEDGE**):

(TRUTH+): If S considers whether p, then
(i) If p is true, S ought to (believe that p)
(ii) If p is false, S ought not to (believe that p) (Hattiangadi 2010: 425).

We have already seen that, according to Bykvist and Hattiangadi (2007), norms like (TRUTH+) will run into difficulties in the case of true blindspot propositions. However, Hattiangadi (2010) assumes for the sake of argument that (TRUTH+) holds, since in her view, this is the most plausible extant truth norm. For unlike (TRUTH*), when p is false and you consider whether p, (TRUTH+) implies that you ought not to believe that p, whereas (TRUTH*) only says that it is not the case that you ought to believe that p.

According to those normativists who are sympathetic to epistemic deontology (e.g. Boghossian 2003), norms construed in terms of obligation and permission are essential to belief. Moreover, the objective norm *explains* why the subjective epistemic norms hold. Although there might be difficulties with *permissive* truth norms (McHugh 2012b; Bykvist and Hattiangadi 2013; Greenberg 201X), it is plausible that among the epistemic norms which purport to be explained by the objective norm is, for instance, (some version of) the following permissive evidential norm:

(PERMISSIVE EVIDENCE): S is permitted to believe that p only if she has sufficient evidence that p.

The problem is that (TRUTH+) cannot explain why we ought to form beliefs in accordance with norms like (PERMISSIVE EVIDENCE) in all cases, because there are instances in which (TRUTH+) and (PERMISSIVE EVIDENCE) clash. I will argue, however, that (KNOWLEDGE**) does not face this difficulty.

Suppose you have excellent evidence for a false proposition *p*. According to (PERMISSIVE EVIDENCE), you are permitted to believe that *p*. But according to (TRUTH+), you ought not to believe that *p*. Since (PERMISSIVE EVIDENCE) and (TRUTH+) yield different permissions and prescriptions (respectively) in this case, the objective norm (TRUTH+) cannot explain and supply the 'rationale' for (PERMISSIVE EVIDENCE) in all cases. Focusing on coin-toss and lottery cases, Hattiangadi (2010) argues that (KNOWLEDGE***) does not clash with (PERMISSIVE EVIDENCE) in this way. But if what I have said is right, then we have no reason to prefer (KNOWLEDGE***) over (KNOWLEDGE**). It remains to be seen that (KNOWLEDGE**) does not clash with (PERMISSIVE EVIDENCE) in the hard cases (i.e. coin-toss and lottery cases). I will argue that it does not.

First, we need to see why – in these hard cases – (TRUTH+) cannot explain (PERMISSIVE EVIDENCE). (I will outline Hattiangadi's (2010: 425-7) argument to show this.) We can begin with coin-toss cases. Suppose you partake in a fair coin-toss, and that you do not yet know the result. In fact, the coin lands on heads. Intuitively, the rational thing to do is to suspend judgement about whether the coin landed on heads until you know the result of the toss. But according to (TRUTH+), you ought to believe that the coin landed on heads, since you consider this proposition and it is true. However, (PERMISSIVE EVIDENCE) implies that you ought to *suspend judgement* about whether the coin landed on heads, since you have no evidence either way. Thus (TRUTH+) and (PERMISSIVE EVIDENCE) yield different prescriptions: the latter, therefore, cannot be explained by the former in all cases. The idea, then, that the objective norm of truth explains why we ought to seek evidential justification for our beliefs faces a difficulty. (That the objective and subjective norms can clash also casts doubt on the idea that the objective norm guides belief formation *indirectly*, via the subjective epistemic norms. I discuss guidance in Chapter 3.) One might agree that following norms like (PERMISSIVE EVIDENCE) *generally* allows believers to satisfy (TRUTH+). The

point is merely that the strategy of explaining (PERMISSIVE EVIDENCE) in terms of (TRUTH+) cannot work in *all* cases.

The same difficulty, however, does not arise for (KNOWLEDGE**). According to (PERMISSIVE EVIDENCE), you ought to suspend judgement about whether the coin landed on heads. But (KNOWLEDGE**) also implies that you ought to suspend judgement. To see this, let p be the proposition "The coin landed on heads". Keeping in mind that, to say that you are in a position to know that p means that, if you were to believe that p, you would know that p, we can argue as follows. If you were to believe that p in this case, (KNOWLEDGE**) implies that you would not know that p – you have no evidential justification for believing that p, or p might be false – and so believing that p violates (KNOWLEDGE**). The best means of conforming to (KNOWLEDGE**) in this case is to suspend judgement about whether p, and this lines up with what, according to (PERMISSIVE EVIDENCE), you ought to do. Likewise, disbelieving that p (i.e., believing that not-p) will not satisfy (KNOWLEDGE**) in this case. If you disbelieve p, it doesn't follow that you are in a position to know that p is false, and so it doesn't follow from (KNOWLEDGE**) that you ought *not* to believe that p. Just as you have no evidential justification for believing that p, you have no justification for disbelieving p either. The best means of conforming to (KNOWLEDGE**) and (PERMISSIVE EVIDENCE), therefore, is to suspend judgement about whether the coin landed on heads. Unlike (TRUTH+), then, (KNOWLEDGE**) gives the right results in the coin-toss case: it prescribes believing what it is intuitively rational to do in this case (i.e., it prescribes suspending judgement).

What about lottery cases? First, let us look at why (TRUTH+) does not give the right results in lottery cases (again following Hattiangadi's presentation of the argument). Lottery propositions are propositions of the form "N's ticket will win". Suppose that you and 99 of your friends each buy a fair ticket, of which there are 100 in total. Your options with respect to satisfying (TRUTH+) are as follows (Hattiangadi 2010: 429):

(a) Believe each lottery proposition.

(b) Disbelieve each lottery proposition.

(c) Suspend judgement about each lottery proposition.

(d) Disbelieve all but one lottery proposition.

According to Hattiangadi, if our aim is to satisfy (TRUTH+), then the best methods of doing so with respect to lottery propositions are option (b) or (d). Option (a) is hopeless, since if only 1 of the 100 tickets will win, then you acquire 99 false beliefs by following (a), and you fail to satisfy (TRUTH+) 99% of the time. Option (b) is more promising, but still problematic. Following (b) means you will acquire 99 true beliefs, and so you will satisfy (TRUTH+) 99% of the time. But you believe that the lottery has a single winner, and if you disbelieve all lottery propositions, you can infer that all the tickets will lose by conjunction-introduction (Douven 2008). This means you will not satisfy (TRUTH+), because one of the lottery propositions is true, and so you ought to believe it. To see the problem, suppose you believe that:

> T_1 *will lose*
>
> T_2 *will lose*
>
> …
>
> T_n *will lose*

It follows that:

> (T_1 *will lose* & T_2 *will lose* & … T_n *will lose*)

However, Hattiangadi argues that this objection to (b) only works if belief is closed under logical consequence. But, according to her, it is not clear that belief is closed under logical consequence. Nevertheless, Hattiangadi argues that there might be further reasons for rejecting (b). For one thing, if (b) is the rational strategy, then (b) implies that it is *always* irrational to buy a lottery ticket no matter the odds. For according to (b), the rational thing to do is to disbelieve *every* lottery proposition – it will be *irrational* to believe that so-and-so's ticket

will win, even if the odds of their winning are extremely high. But suppose you enter a fair lottery with 10 tickets. The prize is 1 million dollars, but the ticket only costs 10 cents. Option (b) implies that it would be irrational to buy this ticket, since according to (b), you epistemically ought to believe that each lottery ticket will lose – buying a ticket that you ought to believe will lose would be irrational according to (b).[59] Intuitively, however, it is not irrational to buy this ticket, since the price of a loss is extremely low and the benefit of a win extremely high. If (b) is the rational strategy, then, it has counterintuitive implications for our judgements about whether buying a lottery ticket is ever rational (Hattiangadi 2010: 429). Moreover, the truth norm purports to explain a *variety* of epistemic norms. Arguably, any plausible system of epistemic norms will include a norm forbidding inconsistent or contradictory beliefs. For example:

(CONSISTENCY*): If p and q are inconsistent, then S ought not to believe that (p & q).

The problem is that if (TRUTH+) also explains (CONSISTENCY*), then (CONSISTENCY*) clashes with (b). For in our lottery case, you believe that only 1 ticket will win; but according to (b), you ought not to believe this – you ought to disbelieve every lottery proposition. And so option (b) violates (CONSISTENCY*). One way to save (b) might be to deny that it is always irrational to hold inconsistent or contradictory beliefs. Consider, for example, one solution to the so-called "preface paradox". Hattiangadi writes:

Authors of books commonly take responsibility, in the preface, for any errors that might remain in the book. Indeed, it seems rational for an author to say this in the preface. However, if you think of a book as a series of sincere assertions, and that the author believes each of the propositions asserted in the book, then if the author expresses the belief that not all of the sentences in the book are true – that at least one of them is false – the

---

[59] Hattiangadi notes that buying the ticket in this case would be *pragmatically* irrational, not *epistemically* irrational. But according to her, this is not to conflate pragmatic and epistemic irrationality: "The point is rather that if [b] is assumed to be the epistemically rational strategy, it will have counterintuitive implications for our judgements regarding whether buying a lottery ticket can ever be practically rational" (2010: 429).

author seems to have expressed inconsistent beliefs. If this is indeed a case in which it is rational to hold inconsistent beliefs, the foregoing explanation for why [b] is not the best method is less compelling. Perhaps there is no norm forbidding contradictory beliefs in the best system of epistemic norms, so [b] is, after all, the best method of belief formation (2010: 429).[60]

But Hattiangadi claims that this would be going too far. Indeed, she argues that – in cases in which it seems rational to hold inconsistent beliefs – we can think of (CONSISTENCY*) as a *prima facie* norm, i.e. a norm that can be overridden by other norms. On this picture, (CONSISTENCY*) might hold in most cases, but it can be overridden in cases such as the preface paradox. As Hattiangadi puts it, "It is less costly to accommodate the preface paradox by treating the non-contradiction norm as being *prima facie* than to deny that the best system of epistemic norms contains a norm forbidding contradictory beliefs" (2010: 429). So, (b) seems to be vindicated.

Option (c) seems intuitively best in the lottery case: your evidence counts against each lottery proposition, and so you ought to suspend judgement about whether your (or anyone's) ticket will win. You ought not to disbelieve each lottery proposition, both because this yields counterintuitive results about practical rationality and because a simply inference will lead you to contradict yourself. So, suspension of judgement is intuitively the best option. But this conflicts with (TRUTH+), since according to (TRUTH+), you ought to believe that so-and-so's ticket will win; it is not the case that you ought to suspend judgement about whose ticket will win.

What about option (d)? As Douven (2008) points out, if you disbelieve all but one lottery proposition, you will at worst acquire 98 true beliefs and at best 100

---

[60] Granted, authors of books don't usually sincerely assert that there *are* some false propositions in their book, just that they take responsibility for any false propositions in the book that there *may* be. But let us simply grant the thought experiment and imagine a bizarre author sincerely asserting that there are some false propositions in the book. Indeed, the preface paradox is especially relevant in the present context, since many claim that any solution to the lottery paradox must yield a solution to the preface paradox, and vice versa (Sutton 2007: 67-70).

of them. So with respect to satisfying (TRUTH+), option (d) is better than the intuitively rational method (c): suspending judgement means that you will not form any beliefs, and so you will not satisfy (TRUTH+), but disbelieving all but one lottery proposition gives you a high chance of satisfying (TRUTH+).

So, according to Hattiangadi, options (b) and (d) are the best means of satisfying (TRUTH+). However, both (b) and (d) conflict with intuitions about the rationality of buying lottery tickets and the rationality of holding contradictory beliefs. Intuitively, the best option is (c): you ought to suspend judgement about what to believe, and this is also what (PERMISSIVE EVIDENCE) tells you to do (you have no evidence about whose ticket will win). Therefore, according to Hattiangadi (2010: 430), the sophisticated explanation of the subjective epistemic norms in terms of the objective norm of truth cannot supply an adequate explanation of intuitively rational norms like (PERMISSIVE EVIDENCE). For according to (PERMISSIVE EVIDENCE), you ought to suspend judgement about what ticket will win; but according to (TRUTH+), you ought to take either option (b) or option (d).

We are now in a good position to see that, unlike (TRUTH+), (KNOWLEDGE**) gives the right results in lottery cases. As in the coin-toss case, both (KNOWLEDGE**) and (PERMISSIVE EVIDENCE) will prescribe suspending judgement about lottery propositions. That is, option (c) is the best means to avoid violating (KNOWLEDGE**) with regard to lottery propositions. Let p be the proposition "N's ticket will win". If you believe that p, you violate (KNOWLEDGE**), since by believing that p, you do not acquire knowledge that p. You have no evidential justification or warrant for believing that p, nor would it be (intuitively) rational for you to believe that p – the lottery is random – and so you are not in a position to know that p. Thus suspension of judgement seems to be the best option, and this lines up with what (PERMISSIVE EVIDENCE) prescribes. Likewise, disbelieving p will also violate (KNOWLEDGE**). For just as in the coin-toss case, from your disbelieving p, it doesn't follow that you are in a position to know

84

that p is false. So if you disbelieve p, it doesn't follow from (KNOWLEDGE**) that you ought not to believe that p. The best means to avoid violating (KNOWLEDGE**) in the lottery case is still option (c), or suspension of judgement. Unlike (TRUTH+) and (PERMISSIVE EVIDENCE), then, (KNOWLEDGE**) and (PERMISSIVE EVIDENCE) do not clash in lottery cases.

2.5 CONCLUSION

This chapter has defended normativism against the objections of Bykvist and Hattiangadi, arguing that a particular knowledge norm does not succumb to those objections. When knowledge is understood as the aim and norm of belief, we preserve the virtues of the truth norm while avoiding its unpalatable consequences: (i) the knowledge norm does not succumb to the objections of Bykvist and Hattiangadi; (ii) it is a plausible candidate for a constitutive doxastic norm; and (iii) it gives the right results in the hard cases (i.e. coin-toss and lottery cases), the upshot of which is that – unlike the truth norm – the knowledge norm does not clash with the plausible epistemic norms.

3 TRUTH, KNOWLEDGE, AND GUIDANCE

3.1 INTRODUCTION

In Chapter 2, we saw that normativism can potentially be defended against the objections Bykvist and Hattiangadi raise for truth norms. This chapter considers a second influential argument against normativism, from Kathrin Glüer and Åsa Wikforss. I want to defend normativism against these objections. In §3.2, I outline the 'no guidance argument' and the related 'regress of motivations' argument from Glüer and Wikforss. §3.3 offers further clarification of the no guidance argument, since it is important to be clear about what the argument amounts to, and a number of commentators have found it puzzling. In §3.4, I suggest one potential way of blocking the objections of Glüer and Wikforss (§3.4.2), following a foray into the literature on 'blind' rule-following (§3.4.1). In §3.5, I outline a different way we might block these anti-normativist arguments, from Teemu Toppinen. I argue that, while we might accept Toppinen's response to Glüer and Wikforss, his own positive account of the way in which norms might guide belief is likely to face significant difficulties: in particular, it relies on an *expressivist* treatment of 'epistemic ought thought'. Although in this chapter I will assume that the truth norm holds – since the issues discussed here are quite independent of the question whether truth, knowledge, or whatever is the proper norm of belief – in §3.6 I briefly sketch a cognitivist account of the way in which the knowledge norm defended in Chapter 2 might guide belief indirectly. Since my positive account will be necessarily incomplete and subject to a number of provisos, I conclude with a sketch of the sort of work that would need to be done in order to make the account convincing.

3.2 THE NO GUIDANCE ARGUMENT

Doxastic normativism is the thesis that it is constitutive of, or essential to, belief that it is norm-governed. According to an influential argument from Kathrin Glüer and Åsa Wikforss ('G&W' hereafter), normativism is problematic because the norms which purport to guide belief formation and revision can offer no such

guidance.[61] Call this *the no guidance argument*. According to G&W, there are two ways one might argue for normativism. On the one hand, one might appeal to a constitutive norm of belief, such as a norm of truth, in order to show that belief is essentially normative. On the other hand, one might appeal directly to the 'rules' of rationality, such as (EVIDENCE). But presumably, if belief is essentially normative, then our belief formation and revision must be guided by the norms in question; that is, it must be the case that we form and revise our beliefs in accordance with them. G&W hold that we cannot do this. I will focus primarily on G&W's objections to the first way of arguing for normativism – the appeal to constitutive norms – since this is the strategy adopted in the preceding chapter.[62]

Now, both the formulation problem (discussed in Chapter 2) and the no guidance argument proceed from highly intuitive assumptions about the prescriptive nature of normative concepts and what it is for a rule or norm to guide belief. Thus it seems that the normativist faces a dilemma:

(a) *Either* she needs to give some novel construal of normativity, in which case she must motivate the departure from intuition.[63]

(b) *Or* she rejects the arguments of both B&H and G&W, in which case her own defense of normativism must proceed from similarly intuitive assumptions.

---

[61] Glüer and Wikforss (2009, 2010a, 2013, 2015). I will focus primarily on the presentation of the no guidance argument in their 2010a, 2013, and 2015 papers, since Glüer and Wikforss acknowledge that their presentation of the no guidance argument in their 2009 is "condensed" and "somewhat obscure" (2010a: 758). This, according to them, is what has puzzled a number of commentators (e.g., Steglich-Petersen 2010, 2013).

[62] Framing G&W's strategy in the way I have done here may invite some confusion, since when outlining their argument against the appeal to truth norms, G&W still talk about 'rule' guidance. In this context, we should understand 'rule' as synonymous with 'norm': to say that you are guided by a rule is just to say that you are guided by a norm. On the other hand, in the context of the second strategy – the appeal to the 'rules' of rationality – the rules in question are just the familiar epistemic norms. G&W have in mind arguments such as Gibbard's (2005), according to which the objective *ought reduces* to the subjective one: the subjective *ought* is analytically prior to the objective one and defines it. Since I will not discuss the latter strategy here, 'rule guidance' should be understood as 'norm guidance'.

[63] G&W suggest that this would amount to accepting their conclusion (2013: 82).

For the norms which purport to be constitutive of belief are typically construed as genuine norms – that is, norms which are prescriptive or action-guiding. According to G&W, as we shall see, this is difficult to square with some basic intuitions about rule- or norm-guidance (2013: 82). In Chapter 2, I opted for strategy (b): the knowledge norm can be formulated, in accord with intuition, in a way that does not succumb to the objections of B&H. But it remains to be seen that the knowledge norm does not succumb to the no guidance problem. I want to suggest, instead, that there may be ways of *blocking* some of G&W's arguments against normativism.

Let us examine first the no guidance argument in the context of the truth norm. Suppose we have the following truth norm for belief:

(TRUTH): *S* ought to believe that *p* if and only if *p* is true.[64]

G&W argue that, given some intuitively plausible assumptions about what rule- or norm-guidance amounts to, we cannot form or revise our beliefs in accordance with (TRUTH). That is, (TRUTH) does not guide our beliefs. Intuitively, according to G&W, to be guided by a rule or norm *R* in our performances requires that *R* "influences, or motivates, or provides reasons for, these performances"; and moreover, "*R* intuitively needs to 'tell us' *what to do under given circumstances*" (2010a: 758. Italics in original). Guiding norms have the following generic form:

(R) Do *X* when in *C*.

Let *C* be the 'antecedent conditions' that must obtain in order to be guided by the norm. Now, the problem is as follows. Some subject *S* looks to *R* for guidance when she is deliberating about whether to *X*. But in order for *S* to figure out whether to *X* on the basis of *R*, she must first form a *belief* about whether the

---

[64] In Chapter 2, we saw that (TRUTH) as it stands is both ambiguous and too strong, and thus normativists typically opt for alternative formulations. The no guidance argument, however, is quite independent of these issues, and so we can assume that (TRUTH) holds for our present purposes. Nor does the no guidance argument challenge the idea that truth norms ever hold in general. Rather, the idea is that truth norms, *if* they hold, would fail to guide belief formation and revision.

antecedent conditions obtain, i.e. she needs to form a belief about whether C is satisfied. If S believes that C is satisfied, then this belief combined with R provides S with a reason to X. For instance:

> [B]eing guided by the rule 'buy low, sell high' requires, among other things, forming a belief about the market. If I believe that the market is at a low, the rule gives me a reason to buy. This belief may of course be false, but this does not prevent the rule from influencing, or motivating, my behavior. Rather, it just means that, if the belief is false, I will (inadvertently) do the wrong thing. Guidance does not necessarily amount to correct performance (G&W 2010a: 758).

This is all well and good, but G&W argue that a problem arises when we apply this picture of guidance to norms like (TRUTH). (TRUTH) is supposed to guide the subject as to whether to believe that p. Whether or not she ought to believe that p is determined by the question of whether p is true or false. But being guided by (TRUTH) requires the subject to form a belief about whether the relevant antecedent conditions obtain. But, in the case of (TRUTH), determining whether the relevant antecedent conditions obtain *amounts to* determining whether p is true, i.e. determining whether C amounts to forming a belief as to whether p. And so, according to G&W, there are two related reasons this makes it intuitively odd to think of (TRUTH) as guiding belief formation and revision. First:

> For one thing, in order to receive guidance as to whether to believe that p from [TRUTH], I must first form a belief as to whether p. But that was the very question I wanted guidance on! Once I have formed a belief as to whether p, I simply do not need such guidance anymore. More precisely, since the very belief the formation of which [TRUTH] is supposed to influence, or motivate, needs to have been formed before [TRUTH] can exert any such influence, [TRUTH] *cannot* influence, or motivate, its formation (G&W 2010a: 758).

Second:

> But it is not only that [TRUTH]'s guidance, so to speak, necessarily comes too late. [TRUTH], secondly, is such that whatever conclusion I come to as to whether p, [TRUTH] 'tells me' that that is the belief I ought to form. That is, whenever I conclude that p is true, [TRUTH] 'gives me a reason' to believe that p. And whenever I conclude that *not-p*, it 'gives me a reason' to believe *not-p*. Hence, [TRUTH] never gives me a reason to believe anything but

what I have already come to believe anyway. Intuitively, no more guidance is to be had from [TRUTH] than from an oracle that invariably tells you to figure out what to do for yourself (G&W 2010a: 759).

The idea, then, is that (TRUTH) cannot guide believers as to whether they should believe *p*, because in order to follow the norm which tells you to believe that *p* on the condition that *p* is true, you need to have already formed a belief as to whether *p* is true. Once you have formed the belief that *p*, there is no longer any work for (TRUTH) to do: it tells you to do what you were doing anyway.[65]

Moreover, doxastic normativism seems to succumb to a version of a related objection G&W raise in their original paper (2009). Call this objection *the regress of motivations* argument.[66] The idea is that not all belief formation can be rule-guided if following or being guided by a rule requires the subject to form a belief about the antecedent conditions, lest we lapse into regress. According to G&W, an intuitive conception of rule-guidance requires a substantive distinction between *being guided* by a rule, and simply *acting in accordance* with one: "For there to be guidance the rule has to make a difference to (the formation and explanation of) *S*'s behavior […] it has to be that *S* acts in accordance with the rule *because* of the rule" (2013: 94). You put one foot in front of the other in order

---

[65] Cf. Shah and Velleman (2005: 519-20), who raise an objection to truth norms that is similar to, but distinct from, the no guidance argument. Unlike G&W, however, Shah and Velleman hold that this objection shows only that the truth norm cannot guide belief *directly*. For a reply to Shah and Velleman, see Zalabardo (2010).

[66] G&W (2009: 55-9; 2010a: 759; 2013: 94-7). In their 2009, G&W discuss the regress of motivations argument in the context of 'content-determining' (CD) normativism (cf. Brandom 1994), as opposed to 'content-engendered' (CE) normativism (cf. Kripke 1982). CD normativism holds, roughly, that there are prior norms in force which *determine* intentional content; CE normativism holds, roughly, that norms can be (directly) *derived* from content. G&W make the CD/CE distinction only in the case of content normativism more generally – in addition to a corresponding distinction in the context of semantic normativism – but it is not clear that the distinction applies in the context of doxastic normativism. Indeed, G&W seem to have abandoned talk of the distinction in their 2010a, 2013, and 2015, which focus solely on doxastic normativism. As such, I will not attempt to apply the distinction in the present context, though see Engel (2007: 184-5) for discussion. Moreover, one might be a doxastic normativist without committing herself to content normativism (and vice versa), so it is not clear that the CD/CE distinction is necessary in present context. Indeed, G&W are clear to distinguish the regress of motivations – which does not rely on a particular theory of content – from what they call the *regress of contents*, which does rely on CD normativism (2009: 57). We need not get into this here.

to walk down the corridor. But this does not mean you are being guided by a rule such as: "When walking down a corridor, put one foot in front of the other!" Putting one foot in front of the other is just something you typically do; you are merely acting in *accordance* with this rule, not following it. So how should we flesh out the relevant notion of rule-guidance? G&W suggest that the most natural, intuitive picture of rule guidance is as follows:

> (RULE GUIDANCE): Some performance A is guided by a rule R if and only if R plays a certain role in S's *motivation* for A. The role R plays in S's motivation for A is such that there will be a certain kind of intentional explanation for A (cf. §3.4 below).

The 'intentional explanation' here involves some kind of *acceptance* of R (G&W 2013: 94; cf. Boghossian 2008: 121-4). G&W claim that, if we use a belief-desire model of intentional explanation, such an explanation will have (at least) two components: (a) acceptance of the rule on the part of S in the 'motivational slot'; and (b) a belief to the effect that some particular performance is in accord with the rule. G&W suggest that upshot of these intuitive ideas about rule-guidance is a *practical syllogism* of the following form:

> (P$_1$) I want to believe what is in accordance with R.
> (P$_2$) To believe that p is in accordance with R.
> (C) I want to believe that p.[67]

G&W argue that this sort of practical inference causes trouble for the normativist:

> The trouble is that the practical inference requires S to have another belief, the belief that believing that p is in accordance with the norm. However, according to the normativist, the further belief would also have to be motivated by a rule, if it is to qualify as a belief, which is to say that there would have to be yet another belief in place. Thus, we embark on a vicious regress, a regress we have called *the regress of motivations*. Hence, if guidance

---

[67] G&W (2009: 55; 2013: 94). See also Glüer and Pagin (1999), who suggest a similar model regarding practical reasoning in the context of rule-following and meaning.

by rules is understood along the lines suggested here, it cannot be that a state is a belief only if it is formed as a result of *S* being guided by certain rules (2013: 94).

According to G&W, therefore, rules or norms cannot guide belief formation, both on pain of regress and because one must have formed the very belief they looked to the norm for guidance on in order to get such guidance.

3.3 THE NO GUIDANCE ARGUMENT: SOME CLARIFICATIONS

It is important to be clear about what the no guidance argument really amounts to. According to Asbjørn Steglich-Petersen (2010: 751), G&W's main claim is the following:

(1) 'Do *X* when in *C*' cannot be guiding if '*X*' and '*C*' refer to the same condition.

But as we shall see in a moment, this is not quite right. Steglich-Petersen gives two objections to (1). First, he argues that G&W are wrong to think that norms with the form of (1) cannot guide. In fact, he says, "If any norm of this form were true, it would be a rather interesting norm with normatively important consequences" (Steglich-Petersen 2010: 751). For example, the norm would entail that one ought to preserve *C* whenever *C* obtains. The norm would be extremely conservative, according to Steglich-Petersen, but it would clearly be capable of providing guidance. Second, Steglich-Petersen claims that it is wrong to think of (TRUTH) as having the form of (1), since norms with the form of (1) are likely to be false: "It is false that believing *p* is correct whenever one believes that *p*" (2010: 752).

The problem, however, is that both of these objections miss the mark. For one thing, G&W claim that it is a mistake to think that norms with the form of (1) can still guide even though they are extremely conservative. For a norm like "If you do *X*, then do *X*" cannot have any consequences for the future (G&W 2010a:

760).[68] Suppose that you ought to *X* whenever you *X*. It does not follow, according to G&W, that you ought to *X* all the time; nor that you ought to continue *X*-ing after you have started.

Moreover – and more importantly – G&W do not claim, contra Steglich-Petersen, that (TRUTH) really has the form of (1). The claim is *not* that (TRUTH) should really be understood as being of the form "If you do *X*, do *X*", since "*p* is true" (*C*) and "believe that *p*" (*X*) are quite obviously distinct (G&W 2010a: 760). Steglich-Petersen notes that (TRUTH) is an objective norm (2010: 752) – a norm that holds in virtue of what is the case – but G&W point out that the fact that it is an objective norm does not necessarily mean that the norm cannot guide.[69] Rather:

> In order to be guided by such norms, one needs to form beliefs as to whether their antecedents are fulfilled, but that does not mean that their antecedents really are 'about' beliefs. Hence, we never claim, as Steglich-Petersen suggests, that what the truth norm 'really recommends is that one should "believe that *p* only if one *believes* that *p*"' [Steglich-Petersen 2010: 752]. Our claim, again, is simply that in order to *determine* whether *C* is fulfilled, *S* already needs to form a belief as to whether *p* – the very belief the formation of which [TRUTH] was supposed to guide (G&W 2010a: 760).

Thus it is wrong to attribute to G&W the view that (TRUTH) really has the form of (1): the claim is not that the antecedent conditions are really 'about' beliefs, but that one must form a belief as to whether *p* in order to *determine* whether *C*. This is thus a significant challenge to the normativist: the normativist needs to

---

[68] Steglich-Petersen is claiming that, on G&W's picture, (TRUTH) really has the form "If you do *X*, then do *X*" because, according to (1), '*X*' and '*C*' refer to the same condition: believing or disbelieving *p*.

[69] Some normativists (e.g. Boghossian 2003) construe objective norms in a way that is likely to invite confusion. They say that objective norms are such that it is not *transparent* to the subject how she might follow them. On the other hand, subjective norms are such that it *is* transparent to the subject how she is to follow them. But this is misleading. Objective norms hold in virtue of what is the case. They do not hold in virtue of anyone's representation of what is the case. So, while it is true that it is not often transparent to subjects how they might follow objective norms, objective norms are not objective norms *because* they are sometimes non-transparent. So, G&W are right to point out that objectivity alone does not prevent belief norms from being guiding norms.

give an account of the way in which the norms essential to belief can guide our belief formation.

Steglich-Petersen's (2010) objection, then, is that the no guidance argument conflates the conditions under which (TRUTH) makes beliefs correct (*p* being true), with the psychological state one must be in in order to apply the norm (the state of believing that *p*). From G&W's (2010a) response, however, it is clear that the no guidance argument involves no such conflation. Conceding this, Steglich-Petersen (2013) makes a second attempt at raising difficulties for the no guidance argument. He argues that the no guidance argument fails because it (I) "relies on a much too narrow understanding of what it takes for a norm to influence behavior" and (II) "falsely assumes that the point of the truth norm is to provide guidance in ascertaining the truth of the propositions being considered for belief" (2013: 279). In their 2015 paper, G&W argue that, once again, Steglich-Petersen's objections miss the mark.

Let us start with (II), since G&W read Steglich-Petersen as claiming that the relationship between (I) and (II) is that, once we recognize the real point of the truth norm, we will be able to see how it can guide belief (G&W 2015: 273). If, as Steglich-Petersen argues, it is false to assume that the point of the truth norm is to guide us in ascertaining the truth of the propositions being considered for belief, then what *is* the point of the truth norm? Steglich-Petersen suggests that we answer *this* question by considering what question the truth norm might provide a helpful answer to: "the truth norm does provide a helpful answer if you were wondering what sorts of considerations would be relevant for determining whether to believe that *p* in the first place" (2013: 283). Recall that the no guidance argument says nothing about whether truth norms are "valid" (i.e., whether they hold or are in force). It only purports to show that truth norms, *if* valid, cannot guide belief formation. G&W (2015: 274) read Steglich-Petersen as suggesting that we need to take a step back and ask not whether truth norms can guide if valid, but *whether they are valid*. According to G&W, Steglich-Petersen

is presupposing that there is a multitude of *possible* or *conceivable* norms for belief, such as:

(PLEASANT): *S* ought to believe that *p* if and only if it would be pleasant to believe that *p*.

If there is such a multitude of belief norms, then we need to ask what the *correctness* of belief depends on, e.g. we need to ask whether it is truth or pleasantness. Steglich-Petersen suggests that the truth norm can provide a helpful answer to this sort of question:

> The truth norm is intended as a guide for what to believe, where this question is initially understood as open to the possibility that something other than truth could decide what it would be correct to believe. It answers this open question by telling us that when considering whether to believe some proposition, one should let this depend on the truth of that proposition […] If one were in doubt as to whether the correctness of belief depended on truth or pleasantness, it would clearly be guiding to become convinced of the truth norm's validity (2013: 281).

On this picture, the point of the truth norm is to answer the question of *which* norms are valid or in force for belief. But G&W note that this is problematic, both because a truth norm can say nothing about its *own* validity, and because even if it could, it would be obscure how it could *convince us* of its own validity: "If the basic question concerning the norms of belief really is which of all the possible norms is in fact valid for belief, it is clearly not any of the norms themselves that will provide us with a helpful answer" (2015: 274-5).

Thus G&W return to (I) above and suggest a second way of interpreting Steglich-Petersen. What Steglich-Petersen is after, according to G&W, is a *wider* or *more inclusive* notion of guidance ("SP-guidance") than theirs. SP-guidance includes the performances which are *excluded* by G&W's narrower notion of guidance, e.g. it includes becoming convinced that the truth norm (rather than the pleasantness norm) is valid for belief (G&W 2015: 275). G&W, by contrast, were originally concerned only with *first-order* belief formation. The no guidance

argument concerned the way in which a truth norm might (not) make a difference in (i.e. *guide*) a subject's forming or revising a first-order belief; it did not involve discussion of some higher-order belief to the effect that the truth norm is valid for first-order believing. Nevertheless, G&W write: "The suggestion we want to investigate now is that Steglich-Petersen's considerations concerning the validity of the truth norm are mere props for investigating the right kind of contrast between different forms of first-order belief formation" (2015: 275).

The relevant 'contrast' in this context is between the outcomes of different *procedures* for determining whether to believe that *p*. One such procedure involves following (TRUTH) and figuring out whether *p* is true; the other involves following (PLEASANT) and figuring out whether *p* is pleasant. It is irrelevant, on this picture, whether a new belief is formed after determining whether the relevant antecedent conditions have been satisfied. What *is* important is whether the *procedure chosen* leads to new beliefs being formed. According to Steglich-Peterson, the chosen procedure does indeed lead to new beliefs being formed: you would end up with a different belief if you followed (TRUTH) than you would if you followed (PLEASANT).

According to G&W, the no guidance argument is part of a larger debate over the 'rule-following considerations' (cf. Boghossian 1989), which bears on the question of whether belief is an essentially rule-guided activity. As noted above, we need to make a substantive distinction between being guided by or following a rule, and merely acting in accordance with the rule. G&W argue, essentially, that (TRUTH) cannot accommodate this distinction: the norm cannot be followed because it cannot guide. The question at issue, then, is whether there is a wider notion of guidance (SP-guidance) which shows the truth norm can be followed.

Steglich-Petersen wants to flesh out a notion of *norm-influenced behavior*, according to which it will plausible that the truth norm can guide or be followed. G&W (2015: 276) reconstruct Steglich-Petersen's picture of norm-influenced behavior as follows:

(D) A norm N of the form "In C, do X" can influence a subject's behavior with respect to X only if S following N can make a difference to S's X-ing (cf. Steglich-Petersen 2013: 281).

Steglich-Petersen fleshes out (D) as follows:

> Given the above comments about the point of the truth norm, it should be clear that it could influence one's behavior, i.e., beliefs, in this sense. For suppose that S instead of accepting [TRUTH] accepts [PLEASANT]. In applying this norm, S will first seek to find out whether believing p would be pleasant. We can imagine that S comes to the conclusion (and thus forms the belief) that believing p would be unpleasant, and, in accordance with the norm, does not form the belief that p. Had S instead followed [TRUTH], she would have sought to find out whether p. If we suppose that she comes to the conclusion that p, she will deem believing p correct. Given that she at this stage in the process *already* believes that p, this will not result in a new belief. But it is nevertheless the case that she would have ended up with a different belief, had she followed [TRUTH] rather than [PLEASANT]. (2013: 282.)

The idea is supposed to be that (D) is weaker than the intuitive notion of guidance G&W rely on, but strong enough that it is nevertheless *sufficient* for rule-following. However, G&W suggest that the scenario Steglich-Petersen gives to flesh out (D) leaves out material he actually uses. Instead, they suggest that Steglich-Petersen has something like the following in mind:

(D*): Acceptance of a norm N of the form "In C, do X" can influence a subject S's behavior with respect to X in the sense of making a difference to S's X-ing only if S can follow N (G&W 2015: 277).

G&W, however, argue that Steglich-Petersen's account (D) – even when construed in its 'causally fortified' version (D*) – gets things exactly backwards (2015: 277). For one thing, it is not clear that Steglich-Petersen's description of (D) establishes that acceptance of (PLEASANT) would influence S's behavior. You accept (PLEASANT), decide that believing p would be unpleasant, and so you do not form the belief that p. This is in *accordance* with (PLEASANT), but it does not follow that S has *followed* the norm. For S might have failed to form the belief

that p for some other reason which is unconnected to her acceptance of (PLEAS-ANT). Thus it is not clear that Steglich-Petersen succeeds in establishing that there is a wider notion of guidance on which (TRUTH) can guide.

Moreover, G&W (2015: 277-8 and fn. 9) suggest that even if we causally fortify (D*) in such a way that we can read 'influence' as implying some belief-inducing causal mechanism which *makes a difference* in subjects' belief formation, it is intuitively clear that, even if this influence is a *necessary* condition on rule-following, it is not *sufficient*. They write:

> This, it seems to us, is just one instance of a point familiar from various attempts at causal analyses of intentional phenomena, most prominently maybe in the theories of action and perception. To be an instance of rule-following, it is not sufficient that acceptance of the rule (causally) influences the relevant performance. It needs to do so *in the right way*. And there just does not seem to be any (informative) way of spelling out what 'the right way' precisely amounts to (2015: 278, fn. 9).[70]

By way of illustration, G&W ask us to imagine the following scenario:

> [A] mad scientist has secretly (and wirelessly) hooked up S's brain to a computer in such a way that if she accepts a norm making the correctness of believing p to depend on property F, considers whether believing p would be F, and comes to the conclusion that it would not, her ability to believe p is (temporarily) blocked. In such a scenario, [PLEASANT], if accepted, influences S's belief formation in the way amounting to SP-guidance – but intuitively, S clearly does not follow [PLEASANT]. Even though accepting the truth norm instead of [PLEASANT] would result in a difference to S's belief formation, the way this difference is generated is no longer recognizable as falling under *any* intuitive notion of rule-guidance. (D*), even in its causally fortified version, gets things backwards. Intuitively, guidance implies influence, but influence can fall short of guidance. While we can derive the possibility of influence from the possibility of guidance, there is no reason to think that the opposite is true. The proper conclusion of these considerations is *not* that our conception of guidance is too narrow but that Steglich-Petersen's is too wide (2015: 278).

If G&W are right, then, norms cannot guide belief: there is no *wider* notion of guidance, such as SP-guidance, according to which they can. The normativist, in

---

[70] I say more about this in §3.4.2 below.

response to these criticisms, would need an account of the way in which norms guide belief that accommodates a substantive distinction between being *guided* by a given rule, and merely acting in accordance with one. In what follows, I attempt to provide such an account.

3.4 BLIND RULE-FOLLOWING

In this section, I want to suggest a potential way of blocking G&W's no guidance argument. G&W, we have seen, maintain that the no guidance argument emerges as one strand of the rule-following considerations. In recent years, the philosophy of language and mind has seen much interest in the notion of *blind* rule-following. A foray into this literature will be necessary in order to illustrate a way in which we might respond to G&W.

In the brief outline of Kripke's Wittgenstein's (KW's) sceptical argument presented in Chapter 1, we saw that one of the grounds on which KW rejects semantic dispositionalism is in its failure to account for the normativity of meaning. KW asked what *makes it the case* that a speaker *means* one thing rather than another by her linguistic expressions, and argued that dispositionalism fails as an answer to this question because (among other things) it leaves out the *oughts* essential to meaning. Dispositionalism is a *reductionist* theory: what a speaker means reduces to the way in which she is disposed to use her expressions – she means *addition* by 'plus' when she is disposed to answer addition-queries with the *sum* of the numbers added, rather than answer according to some other function (such as the 'quus' function). By contrast, *non-reductionism* holds (roughly) that meaning is of a more primitive character: it is *sui generis* and cannot be characterized in non-semantic and non-intentional terms.

KW is not happy with non-reductionism as a response to the sceptic. Suppose we want to say that meaning is a primitive state, *sui generis* and non-reducible. KW gives the following objection to a move towards such a view:

> Such a move may in a sense be irrefutable, and if it is taken in an appropriate way Wittgenstein may even accept it. But it seems desperate: it leaves the nature of this postulated

primitive state – the primitive state of 'meaning addition by "plus"' – completely mysteri-
ous. It is not supposed to be an introspectible state, yet we supposedly are aware of it with
some fair degree of certainty whenever it occurs. For how else can each of us be confident
that he *does*, at present, mean addition by 'plus'? (Kripke 1982: 51).

A number of philosophers have found this objection to non-reductionism inade-
quate. For one thing, KW simply leaves it unargued that non-reductionism is
desperate and mysterious, as if it were obvious. Here is Colin McGinn:

[KW's point is that] once we abandon the idea that meaning is an irreducible *experiential*
state we have no account of the nature of our first-person knowledge of meaning – we have
no conception of how the non-experiential primitive state is an object of distinctively first-
person knowledge […] How to give a philosophical *theory* of this kind of knowledge is of
course a difficult and substantive question, but the lack of a theory of a phenomenon is
not in itself a good reason to doubt the *existence* of the phenomenon. I therefore see no
mystery-mongering in the claim that there are primitive non-experiential mental states
which display a distinctive first-person epistemology (McGinn 1984: 160-61).[71]

Likewise, in an influential survey of the literature on KW, Paul Boghossian
(1989) argues that non-reductionism survives KW's objections (though he does
concede that KW's objections to *reductionist* theories (such as dispositionalism)
warrant rejection of those theories). KW's charge is that we would not be able to
explain our thoughts if we endorsed a non-reductionist account of mental con-
tent. Boghossian notes that even though the classic problems of self-knowledge
are extremely difficult, in order for KW's objection to non-reductionism to be
convincing, we would need a *proof* that we could not develop a satisfactory first-
person epistemology. Boghossian writes:

Kripke, however, provides no such proof. He merely notes that the non-phenomenal char-
acter of contentful states precludes an introspective account of their epistemology. And
this is problematic for two reasons. First, because there may be non-introspective accounts
of self-knowledge. And second, because it does not obviously follow from the fact that a
mental state lacks an individuative phenomenology, that it is not introspectible […] In the
context of Kripke's dialectic, the anti-reductionist suggestion emerges as a stable response

---

[71] See Wright (2001: 113 ff.) for an objection to McGinn and further discussion.

to the sceptical conclusion, one that is seemingly untouched by all the considerations adduced in the latter's favor (1989: 42-3).

Nevertheless, from the fact that Kripke's objection to non-reductionism is inadequate it does not follow that non-reductionism is true. Indeed, Boghossian notes that there are a number of significant difficulties with non-reductionism, so that even if we endorse a non-reductionist position as a response to Kripke's sceptic, much work still needs to be done (1989: 48-50). Boghossian defines *robust realism* about meaning, which includes non-reductionism, as the following view:

> (ROBUST REALISM): Judgements about meaning are factual, irreducible, and judgement-independent.

Now, according to Boghossian, even if robust realism/non-reductionism survives KW's objections, it faces at least three significant difficulties of its own:

> Robust realism harbors some unanswered questions, the solutions to which appear not to be trivial. There are three main difficulties. First: what sort of room is left for *theorizing* about meaning, if reductionist programs are eschewed? Second: how are we to reconcile an anti-reductionism about meaning properties with a satisfying conception of their causal or explanatory efficacy? And, finally: how are we to explain our (first-person) knowledge of them? (1989: 48).[72]

More recently, however, Boghossian has argued that even if these challenges to non-reductionism can be met in such a way that we can successfully respond to KW's sceptical challenge, there is still a *more fundamental* problem with non-reductionist conceptions of rule-following (Boghossian 2008, 2012).[73] Let us look at this (alleged) more fundamental problem.

---

[72] We need not go into the details of these difficulties with non-reductionism – I want only to note that there are some non-trivial ones. I will say more below about how Boghossian's appeal to a non-reductionist conception of content might diffuse KW's sceptical challenge.

[73] Wright (2007, 2012) endorses a similar view. To keep things manageable, I will focus only on Boghossian's (more recent) arguments.

3.4.1 THE INFERENCE PROBLEM AND A SOLUTION

According to Boghossian, the most pre-theoretic and intuitive model of rule-following involves four related components (2008: 119; 2012: 32):

(ACCEPTANCE): If S is following a given rule R ('If C, do A'), then S has somehow accepted R.

(CORRECTNESS): If S is following rule R, then S acts correctly *relative to his acceptance* if it is the case that C and he does A.

(EXPLANATION): If S is following rule R by doing A, then S's acceptance of R *explains* S's doing A.

(RATIONALIZATION): If S is following rule R by doing A, then S's acceptance of R rationalizes S's doing A.

Following Crispin Wright (2012: 383), call this the 'ACRE' model of rule-following. Now, suppose that S is following the rule:

(EMAIL): Answer any email that calls for an answer immediately upon receipt!

Boghossian writes:

> [I]t is clear that it would be correct to say that I was following [EMAIL] in replying to the email, rather than just coincidentally conforming to it, when it is somehow or other *because* of [EMAIL] that I reply immediately […] [F]or me to be following the rule, the 'because' must be that of rational action explanation: I follow [EMAIL] when that rule serves as *my reason* for replying immediately, when that rule *rationalizes* my behavior. I want to suggest, then, that the minimal content of saying that person S follows rule R in doing A is that *R serves as S's reason for doing A* (2012: 31).

On our ACRE model, then, if S is following (EMAIL), then S has *accepted* (EMAIL) rather than, for instance, some alternative rule:

(EMAIL*) Answer any email that calls for an answer, unless it's from your supervisor, immediately upon receipt!

Moreover, it is *correct*, relative to S's acceptance of (EMAIL) rather than (EMAIL*), for S to answer *any* email that calls for an answer immediately upon receipt (even if the email is from S's supervisor). If S accepted (EMAIL*) rather than (EMAIL), then it would be correct, relative to her acceptance of (EMAIL*), to ignore the email from her supervisor and answer all others that call for an answer immediately upon receipt. Further, S's acceptance of (EMAIL) *explains* why she answered immediately the email she just received – she answered the email *because* of the rule, where 'because' denotes rational action explanation. And finally, (EMAIL) *rationalizes* (i.e. helps us to understand and make sense of) S's immediately answering the email she just received.

Note that the ACRE model is primarily about *personal-level* rule-following. For *sub-personal* rule-following, we remove (RATIONALIZATION): "If I say of a calculator that it is adding, then I am saying that its 'internalization' of the rule for addition (via programming) explains why it gives the answers that it gives. But I am obviously not saying that the addition rule *rationalizes* the calculator's answers. The calculator doesn't act for reasons, much less general ones" (Boghossian 2012: 32).

KW's discussion is primarily about personal-level rule following, and on (ACCEPTANCE) as a condition for it. KW's sceptical challenge, recall, is this: what fact about me is constitutive of my meaning one thing rather than another by some linguistic expression? To use Kripke's example, suppose we have two rules:

(ADDITION): When queried about the answer to the question '$x + y = ?$' answer with the *sum* of x and y!

And:

(QUADDITION): When queried about the answer to the question '$x + y = ?$' answer with the *quum* of x and y!

What fact makes it the case that I am following (ADDITION) rather than (QUADDITION)? Likewise, we might ask what makes it the case that I am following

(EMAIL) rather than (EMAIL*). This is the sceptical challenge concerning (AC-CEPTANCE).

Now, we have seen that Boghossian (1989) argues we can meet the sceptical challenge by appeal to a non-reductionist conception of content, despite the difficulties with that conception. Boghossian (2008, 2012) now argues that even if this is right, there is a more fundamental problem about rule-following. Consider the following view, which Boghossian attributes to Wright (2001):

> (INTENTION VIEW): We follow rules by forming intentions to uphold certain patterns in our thought or behavior and by acting on those intentions (Boghossian 2012: 27).[74]

Boghossian holds that we can diffuse KW's sceptical worry if we combine (IN-TENTION VIEW) with the *rejection* of the following view:

> (MEANING ASSUMPTION): Expressions get their meaning by our following rules in respect of them (Boghossian 2012: 36).

How does (INTENTION VIEW) combined with the rejection of (MEANING ASSUMPTION) purport to dissolve KW's sceptical worry?[75] We must give up either (IN-TENTION VIEW) or (MEANING ASSUMPTION) because combining both views yields the consequence that rule-following – and with it mental content – is *metaphysically impossible*:

> [G]iven the two assumptions [i.e. (INTENTION VIEW) and (MEANING ASSUMPTION)], we would be able to reason as follows. In order to follow rules, we would antecedently have to have intentions. To have intentions, the expressions of our language of thought would have to have meaning. For those expressions to have meaning, we would have to use them according to rules. For us to use them according to rules, we would antecedently have to have

---

[74] According to Boghossian, (INTENTION VIEW) is really a special instance of a more general view, "according to which rule acceptance consists in some intentional state or other, even if it is not identified specifically with an intention" (2012: 33). This Boghossian calls the *intentional* view.

[75] Boghossian adds a number of provisos to the idea that we can meet KW's sceptical challenge by appeal to non-reductionism, the details of which need not concern us here (2012: 35 ff).

intentions. And so neither content nor rule-following would be able to get off the ground (Boghossian 2012: 36).

So it looks like we must reject either (INTENTION VIEW) or (MEANING ASSUMPTION), lest we be forced to concede that rule-following (and therefore content) is metaphysically impossible. KW rejects (INTENTION VIEW), since according to Boghossian, KW regards (MEANING ASSUMPTION) as *non-optional* (2012: 36). Boghossian himself, however, thinks we should reject (MEANING ASSUMPTION), and this is what leaves non-reductionism an open option as an answer to the sceptic. If we hold onto non-reductionism about content, then we appear free to endorse (INTENTION VIEW), since we can construe meaning as primitive, *sui generis*, and not explicable in non-semantic and non-intentional terms.

But Boghossian now argues that even assuming the (significant) challenges to non-reductionism can be met in such a way that we can successfully respond to KW's sceptic, there is still a *more fundamental* problem about rule-following. What is this more fundamental problem? (INTENTION VIEW), according to Boghossian, suffers from "a further and seemingly fatal flaw" (2012: 39) which concerns (EXPLANATION) and (RATIONALIZATION), rather than (ACCEPTANCE).

To see the problem, suppose I have adopted the (EMAIL), where this adoption consists in my forming an explicit intention to conform to the instructions of the (EMAIL) rule:

(INTENTION): For all *e*, if *e* is an email and you have just received *e*, answer it immediately!

But how should we understand my being guided by these instructions, i.e. my following the (EMAIL) rule? Boghossian writes:

To act on this intention, it would seem, I am going to have to think, even if very fleetingly and not very consciously, that its antecedent is satisfied. The rule itself, after all, has a conditional content. It doesn't call on me to just do something, but to always perform some action, if I am in a particular kind of circumstance. And it is very hard to see how such a

conditional intention could guide my action without my coming to have the belief that its antecedent is satisfied (2012: 40).

That is, in order to follow the (EMAIL) rule, I will need to think to myself (i.e. *believe* that):

(BELIEF): *e* is an email I have just received.

In order to reach the following conclusion, which (following Miller 2015) we might call an *action-guiding state* (AGS) with the content:

(AGS) Answer *e* immediately!

That is, in order to follow (EMAIL), we need (INTENTION) to uphold the instructions of the rule ("For all *e*, if *e* is an email and you have just received *e*, answer *e* immediately!") combined with (BELIEF) ("*e* is an email I have just received") in order to arrive at (AGS), a psychological state in which you are motivated to act (i.e. a state with the content: "Answer *e* immediately!").[76] According to Boghossian, in this instance, rule-following on the (INTENTION VIEW) requires *inference*: "it requires the rule-follower to infer what the rule calls for in the circumstance in which he finds himself" (2012: 40). Call this Boghossian's *inference problem*.

However, Boghossian maintains that the (EMAIL) rule is hardly special: "Since *any* rule has a *general* content, if our acceptance of a rule is pictured as involving its representation by a mental state of ours, an inference will always be required to determine what action the rule calls for in any particular circumstance" (2012: 40). The upshot is that, on the (INTENTION VIEW), following a rule will *always* involve inference on the part of the subject.

To make the problem clear, consider another example. Consider an *imperatival* equivalent to the *modus ponens* rule:

---

[76] As we did in previous chapters, we can set aside any Humean worries about belief, desire, motivation, and the relationship between them.

(MP*): From 'If *C*, do *A*' and *C*, conclude 'do A'!

Boghossian's claim is that the inference problem proceeds from a pattern of reasoning which – again following Wright (2007: 491) – we might call the *modus ponens model*. Following (EMAIL) is like following (MP*). Boghossian writes:

> On the Intention View, applying the Email Rule requires, as we have seen, having an intention with the rule as its content and inferring from it a certain course of action. However, inference, we have said, involves following a rule, in this case, MP*. Now, if the Intention View is correct, then following the rule MP* itself requires having an intention with MP* as its content and inferring from it a certain course of action. And now we would be off on a vicious regress: inference rules whose operation cannot be captured by the intention-based model are presupposed by that model itself (2012: 41).

Boghossian's main contention is that the following five propositions form an inconsistent set (2012: 42):

(1) Rule-following is possible.

(2) Following a rule consists in acting on one's acceptance (or 'internalization') of a rule.

(3) Accepting a rule consists in an intentional state with general (prescriptive or normative) content.

(4) Acting under particular circumstances on an intentional state with general (prescriptive or normative) content involves some sort of inference to what the content calls for under the circumstances.

(5) Inference involves following a rule.

If Boghossian is right, then we must give up on one of these claims. But which one? Denying (1) would amount to scepticism about rule-following of the sort KW's sceptic endorses. (2) is true if our ACRE model is indeed the most pre-theoretic and intuitive picture of rule-following – it is the "minimal content of saying that someone is following a rule" (Boghossian 2012: 42). (3) is just the general intentional view, of which (INTENTION VIEW) is one particular instance. And (4) also seems true: consider what it takes to follow (EMAIL) – we need an

107

inference from (for example) the rule and (BELIEF) in order to reach (AGS). Finally, (5) seems obvious: inference is an example of rule following *par excellence*, and so Boghossian maintains that (5) is "analytic of the very idea of deductive inference" (Boghossian (2012: 42).[77]

Now, according to Boghossian, the only plausible option seems to be to give up on (3) – the intentional view, and with it (INTENTION VIEW). Thus in order to avoid the conclusion that rule-following is impossible, "we must find a way of understanding the notion of *accepting* or *internalizing* a rule that does not consist in our having some *intentional* state in which that rule's requirements are explicitly represented" (Boghossian 2012: 42).[78] Boghossian suggests that Wittgenstein arrived at the same conclusion in the *Philosophical Investigations*:

> "All the steps are really already taken" means: I no longer have any choice. The rule, once stamped with a particular meaning, traces the lines along which it is to be followed through the whole of space. – But if something of this sort really were the case, how would it help me?
>
> No; my description made sense only if it was to be understood symbolically. – I should say: *this is how it strikes me*.
>
> When I follow the rule, I do not choose.
>
> I follow the rule *blindly* (1953: §219).

The upshot, according to Boghossian, is that some rule-following must be *blind*: "in its most fundamental incarnation, rule acceptance cannot consist in the for-

---

[77] G&W disagree: "it is not clear to us at all that inference involves following a rule. What Boghossian's argument shows, at most, is that inference is yet another controversial case creating trouble for any substantive understanding of rule-following" (2010b: 163). G&W suggest rejecting (5), since they do not share the intuition that inference always involves following a rule. According to them, one reason we might deny that (5) is a conceptual truth is that it may have unpalatable consequences, such as the consequence that rule-following is impossible. But as we shall see below, we need not abandon the idea that rule-following involves following a rule (5), nor the intentional view (3). Rather, we may only need to make a slight modification to (4).

[78] For an alternative, purportedly Fregean account of non-intentional rule-following, see §1.12 and §3.15 of Ichikawa and Jarvis (2013). There are likely to be highly contentious issues with this, and I cannot properly explore them here.

mation of a propositional attitude in which the requirements of the rule are explicitly encoded" (2012: 43). It is not an option to embrace rule-following scepticism, Boghossian suggests, since this seems false about reasoning in general and about the intelligibility of deductive inference: the reasons a subject has for moving from certain premises to certain conclusions are in general of the *essence* of deductive inference. But this brings us into an 'antinomy of pure reason':

> So what we are contemplating, when we contemplate giving up on the Rulish picture of deductive inference, is not so much giving up on a Rulish *construal* of deductive inference as giving up on deductive inference itself. But that is surely not a stable resting point – didn't we arrive at the present conclusion through the application of several instances of deductive inference? Hence we have what I have somewhat grandly called an "antinomy of pure reason": we both must – and cannot – make sense of the notion of someone's following a rule. The only non-skeptical option that seems open to us is to try taking the notion of following – or applying – a rule as primitive, effectively a rejection of proposition (4) above (Boghossian 2012: 47).

Note that Boghossian's ultimate conclusion goes *beyond* the anti-reductionism about content (endorsed in Boghossian 1989) as a response to KW's sceptic. The present view involves primitivism about rule-following or rule-application *itself*: "we would have to take as primitive a *general (often conditional) content serving as the reason for which one believes something*, without this being mediated by inference of any kind" (2012: 47).

Thus if Boghossian is right, then even if the difficulties with non-reductionism can be met in order to successfully respond to KW's sceptic, then there is a *further* problem – the inference problem – for rule-following. This problem would seem to require us to abandon the intentional view, and with it (INTENTION VIEW), lest we give up either some highly intuitive assumptions about rule-following, or the idea that rule-following is so much as possible. In response to Boghossian (and Wright 2012), however, Alex Miller (2015) argues that there is no *further* problem about rule-following of the sort Boghossian raises. Miller's response, I want to suggest, may allow us to block some of G&W's arguments

against normativism. First, we need to look at Miller's solution to the inference problem.

Boghossian's 'antinomy of pure reason' was that we *must*, and at the same time *cannot*, make sense of someone's following a rule. Miller argues that we can avoid the antinomy by appeal to Wittgenstein's idea that, in order to avoid certain paradoxical implications of rule-following, we need to abandon the idea that following a rule is a matter of *interpretation*. As Wittgenstein (somewhat cryptically) puts it:

> This was our paradox: no course of action could be determined by a rule, because every course of action can be brought into accord with the rule. The answer was: if every course of action can be brought into accord with the rule, then it can also be brought into conflict with it. And so there would be neither accord nor conflict here.

> That there is a misunderstanding here is shown by the mere fact that in this chain of reasoning we place one interpretation behind another, as if each one contented us at least for a moment, until we thought of yet another lying behind it. For what we thereby show is that there is a way of grasping a rule which is *not* an interpretation, but which, from case to case of application, is exhibited in what we call "following a rule" and "going against it" (1953: §201).

The idea is that, if we hold on to the idea that rule-following requires interpretation, then in order to apply rule R in a given case, we will need to apply a further rule R^* in order to determine what R amounts to in the given case. This seems to have the paradoxical upshot that rule-following is impossible. For under the idea that rule-following requires interpretation, we seem to get an infinite regress:

> Applying rule R in a given case on the basis of interpretation requires the application of rule R^* in order to determine what R amounts to in this case.

> Applying rule R^* in a given case on the basis of interpretation requires the application of rule R^{**} in order to determine what R^* amounts to in this case.

Applying rule R^{**} in a given case on the basis of interpretation requires the application of rule R^{***} in order to determine what R^{**} amounts to in this case.

…

And so on *ad infinitum*. Wittgenstein's solution to this problem is to reject the idea that rule-following always requires interpretation. Miller writes:

> Against the idea that rule-following necessarily involves interpretation, Wittgenstein's suggestion seems to be that in applying a rule R in a particular case there need be no *further* inferential step – over and above that involving R *itself* – mediating between acceptance of R and that particular application. As Wittgenstein puts it in §228 of *Philosophical Investigations*, '[W]e look to the rule for instruction and *do something*, without appealing to anything else for guidance' (Miller 2015: 405).

How might this help to deal with the inference problem? Miller's argument is that, once we abandon the idea that rule-following requires interpretation, we need not apply MP* in order to apply (EMAIL): we need only to apply (EMAIL). According to Miller:

> When we say that there is a way of following [EMAIL] that is not an interpretation we are saying that between acceptance and application of the [EMAIL] rule there is no additional application of a further rule (such as MP*). We might say: we look to the [EMAIL] rule for instruction and so something – reply to the email – without appealing to anything else for guidance, *a fortiori* without appealing to MP* for guidance. Following Wittgenstein's lead by rejecting the idea that rule-following always involves interpretation thus stops Boghossian's regress in its tracks (2015: 406).

As per Wittgenstein's suggestion, there need not be any *further* inferential step, "over and above" that involving the (EMAIL) rule itself, which mediates between one's acceptance of the (EMAIL) rule and a particular application of it. Upon receiving some email e that calls for an answer, we look to (EMAIL) for instruction and 'do something' – reply to the email. We need not appeal to a further rule, such as MP*, in order to be guided by (EMAIL). If Miller is right, then, to give up

on the idea that rule-following always requires interpretation is to block Boghossian's inference problem.

But what does this mean for the ACRE model? Miller argues that, if we reject the conception of rule-following as interpretation, we can still preserve each component of the ACRE model, as well as propositions (2)-(5) of Boghossian's inconsistent set (though with a slight modification). First, Miller claims that there is no inference, based on the application of a further rule (such as MP*), which mediates between acceptance and application of the (EMAIL) rule. What, then, *does* mediate between acceptance and application of the rule? Recall that, on Boghossian's picture, we first need an *intention* to comply with the instructions of the (EMAIL) rule, combined with a belief to the effect that the antecedent of the conditional rule is satisfied, in order to reach a psychological state the content of which complies with the instructions of the rule. That is, we first need an intention to comply with the instructions:

(INTENTION): For all *e*, if *e* is an email and you have just received *e*, answer it immediately!

And we combine (INTENTION) with the following *belief*:

(BELIEF): *e* is an email I have just received.

In order to reach an action-guiding state with the content:

(AGS): Answer *e* immediately!

But what, Miller asks, mediates the "causal transition" from (INTENTION) and (BELIEF) to (AGS), when we give up on the conception of rule-following as interpretation (2015: 407)? Miller appeals to Wittgenstein's answer to questions of this sort. Wittgenstein's well-known answer to questions of this sort appeals to the ideas of *training* and *initiation into a practice or custom*. Take two representative passages from Wittgenstein's *Philosophical Investigations*:

> Following a rule is analogous to obeying an order. One is trained to do so, and one reacts to an order in a particular way (1953: §206).

> The absent-minded man who at the order "Right turn!" turns left, and then, clutching his forehead, says "Oh! Right turn", and does a right turn. – What has struck him? An interpretation? (1953: §506).

When viewed in light of the Wittgensteinian ideas of training and initiation into a practice or custom, we have an account of the mediation between acceptance of a given rule and a given application of that rule. On the present example, Miller writes:

> What happens in the [EMAIL] case is that the agent S is in an intentional state [INTENTION] with the content *If you receive an email, answer it!* which together with his belief *that e is an email* (in the right kind of way) causes S to reply to the email (or more precisely to have an action-guiding state with the content *answer e!*). This causal relation is set up and sustained by facts about custom, practice and training (2015: 407).[79]

Now, we are in a position to see the way in which Miller's account allows us to preserve Boghossian's propositions (2)-(5), though with one slight modification. Boghossian's second proposition was this:

(2) Following a rule consists in acting on one's acceptance (or internalization) of a rule.

According to Miller, (2) can be dealt with in a straightforward way, since – courtesy of a non-reductionist conception of content – we can help ourselves to the notion of intentional content. In following (EMAIL), the (BELIEF) (that *e is an email I just received*) and (INTENTION) cause the subject to be in the action-guiding state: "following rule R is a matter of being caused (in the right kind of way)

---

[79] Miller notes that the point is not that we want to recover the intentional content of the relevant states (i.e. the intention, the belief, and the action-guiding state) in terms of their causal relations to other states. Rather, the point is that – *assuming* non-reductionism about content allows us to secure the notion of intentional content against KW's sceptical challenge – there is no *further* problem (such as Boghossian's inference problem) concerning rule-following. Moreover, Miller uses 'in the right kind of way' in order to exclude deviant causal chains. This is important, but I want to postpone discussion of it until §3.4.2 below.

to act in a way that conforms to R inter alia by the having of an intentional state with the rule R as its content" (Miller 2015: 408). Proposition (3) is also straightforward:

> (3) Accepting a rule consists in an intentional state with a general (prescriptive or normative) content.

Proposition (3) can be dealt with straightforwardly because we can say that "accepting the [EMAIL] rule consists in being in an intentional state with the content *If you receive an email, answer it!*" (Miller 2015: 408). Since the content of the intentional state ("If you receive an email, answer it!") is imperatival in form, it should be clear that the content of the state is indeed prescriptive or normative in character.

Now, Miller argues that proposition (4) needs a slight modification. Boghossian's proposition (4) was this:

> (4) Acting under particular circumstances on an intentional state with a general (prescriptive or normative) content involves some sort of inference to what the content calls for under the circumstances.

Given our foregoing Wittgensteinian considerations, Miller suggests we modify this as follows:

> (4*) Acting under particular circumstances on an intentional state with a general (prescriptive or normative) content involves some sort of *rule-mediated transition* to what the content calls for under the circumstances.

Thus according to (4*), the (EMAIL) rule is what mediates the transition from the belief (i.e. the state with the content: "*e* is an email") to the action-guiding state (i.e. the state with the content: "Answer *e*!"). In this instance, S follows (EMAIL) rather than MP*. Granted, the transition between the belief and the action-guiding state *conforms* to MP*, but S is not required to *follow* MP* in order to bring

about the action-guiding state via (INTENTION) and (BELIEF) (Miller 2015: 408).[80] This is noteworthy, since as we have already seen, any plausible account of rule-following must accommodate a substantive distinction between *following* or *being guided by* a rule, and merely *conforming to* or *acting in accordance* with one (more on this below).

Finally, what of proposition (5)? This was:

(5) Inference involves following a rule.

There should be nothing preventing us from accepting (5) if we can view "inference qua psychological process in terms of being caused (in the right kind of way) to draw a conclusion by an intention to uphold a rule of inference like MP*" (Miller 2015: 408). Moreover, as Boghossian points out, inference is an example of rule-following *par excellence*, so (5) looks intuitively plausible.

Now, as for our ACRE model, Miller points out that there is no problem with (ACCEPTANCE) and (CORRECTNESS): "My following the [EMAIL] rule certainly involves my acceptance of the rule, and if I reply to a recently received email I act correctly relative to my acceptance of the rule, and incorrectly otherwise" (2015: 408-9). Miller's account also seems to capture (EXPLANATION) and (RATIONALIZATION). For acceptance of the (EMAIL) rule, for example, consists in an intentional state (INTENTION) "from which there is a causal route (of the right kind) to my having an action-guiding state with the content *answer e!*" (Miller 2015: 409). This is a *causal explanation*, and so it captures (EXPLANATION) and, it seems, (RATIONALIZATION). It captures (RATIONALIZATION) because our causal story about the movement from (INTENTION), to the belief that the antecedent of the rule is satisfied, to the action-guiding state *rationalizes* (i.e. makes sense of) my answering the recently received email.

---

[80] As Miller points out, it follows that what Wright calls 'the *modus ponens* model' is something of a misnomer.

Thus it looks like Boghossian's inference problem can be met. If Miller is right, then – courtesy of a non-reductionist conception of content – there is no *further* problem about rule-following. But does this mean we need to reject the idea of *blind* rule-following? In one sense, yes; in another, no. For there seem to be two ways in which we might understand the notion of blind rule-following. The first way of understanding blind rule-following – which, if Miller is right, we can reject – is Boghossian's interpretation of Wittgenstein:

> [W]e have shown that, in its most fundamental incarnation, rule-acceptance cannot consist in the formation of a propositional attitude in which the requirements of the rule are explicitly encoded. Such a picture would be one according to which rule-following is always fully *sighted*, always fully informed by some recognition of the requirements of the rule being followed. And the point that Wittgenstein seems to be making [at §219 of *Philosophical Investigations,* quoted above] is that, in its most fundamental incarnation, not all rule-following can be like that – some rule-following must simply be *blind* (Boghossian 2008: 130).

According to Boghossian's reading of Wittgenstein, then, the idea that rule-following is blind means we must reject the intentional view, and with it (INTENTION VIEW). This lands us in the 'antinomy of pure reason' when we combine rejection of the (INTENTION VIEW) with KW's arguments against the possibility of non-intentional rule-following. But Miller argues that Boghossian's reading of Wittgenstein is mistaken:

> [Boghossian's] reading of Wittgenstein's use of 'blind', however, seems to me to be mistaken. When Wittgenstein speaks of rule-following as 'blind', 'blind' does not contrast with 'intentional' (as in Boghossian's reading). Wittgenstein is not arguing that acceptance of a rule is not an intentional state. Rather, 'blind' contrasts with 'based on an interpretation'. To say that following a rule is ultimately blind is to say that at the fundamental level following a rule is not based on interpretation (2015: 411).[81]

To see this, consider again our (EMAIL) rule. Suppose that *S*'s application of (EMAIL) *is* based on interpretation, and she is asked for a reason or justification

---

[81] This reading of Wittgenstein is shared by John McDowell (1998: Chapters 11 and 12). See also Wright (2001: Chapters 5 and 6) for responses to McDowell and exegesis of Wittgenstein.

for why, upon receiving an email, she applied it in the way she did. In order to supply such a justification for her application of (EMAIL), *S* would need to cite a *further* rule over and above (EMAIL) – perhaps she cites MP*, combined with her belief to the effect that the antecedent of the rule is satisfied ("*e* is an email I have just received") in order to explain why she answered the email. But this is not what happens when we abandon the idea that rule-following requires interpretation:

> [S]ince my following the [EMAIL] rule is not based on an interpretation, when I'm asked for a justification or reason for why I applied it in the way that I did in that situation I can tell no such story: all I can say is *This is how it strikes me* (or, with Wittgenstein [1953]: §217, *This is simply what I do*). In the passage in §219 of *Philosophical Investigations*, therefore, Wittgenstein is not raising a difficulty for the construal of rule-acceptance as an intentional state: rather, he is reiterating the need to reject the conception of rule-following as interpretation if we are to avoid the conclusion that there are no determinate facts about which rule an agent is following (Miller 2015: 412).

If Miller is right, then courtesy of a non-reductionist conception of content, there is no further problem concerning rule-following; we can appeal to the notions of practice/custom/training and blind rule-following (where 'blind' contrasts with 'based on interpretation') in order to preserve our ACRE model of rule-following and (INTENTION VIEW). I want to suggest that this provides us with the tools for responding to Glüer and Wikforss.

3.4.2 REPLY TO GLÜER AND WIKFORSS

Given our foray into recent literature on the rule-following considerations, we are now in a position to sketch a way in which an appeal to blind rule-following (where 'blind' contrasts with 'based on interpretation') might enable us to block G&W's arguments against normativism. To begin with, note that there are structural similarities between both G&W's no guidance argument and the regress of motivations argument, and Boghossian's inference problem. Consider a rule of the form:

(R) Do *X* when in *C*

According to the no guidance argument, norms or rules cannot guide belief, because in order to determine whether C, the subject needs to form a *belief* as to whether C. That is, in order to determine whether to believe that p on the basis of a norm like (TRUTH), you need to form a belief as to whether the antecedent conditions of the norm are satisfied, i.e. as to whether p. But since the question of whether to believe that p is precisely the question on which the subject looks to (TRUTH) for guidance, the norm can offer no such guidance, because the subject needs to form the very belief she wanted guidance on in order to be guided by that norm.

Lurking in the no guidance argument is the regress of motivations problem, according to which not all belief formation can be rule guided, because if it were, it would set us on a vicious infinite regress. Consider again the following practical syllogism:

(P₁) I want to believe what is in accordance with R.

(P₂) To believe that p is in accordance with R.

(C) I want to believe that p.

The problem was that (P₂) is a second-order belief (i.e. a belief about a belief) to the effect that believing that p is in accordance with R. If all belief is rule- or norm-guided, then we must appeal to a *further* rule R^* to guide the second-order belief (P₂), and yet another belief about whether the antecedent of R^* is satisfied, in order to be motivated to believe that p. But then we must appeal to *yet another* rule R^{**}, then another R^{***} and so on *ad infinitum*. We land on a vicious infinite regress.

Boghossian's inference problem is similar in spirit to the regress of motivations argument. According to Boghossian, we have seen, the move from (INTENTION) and (BELIEF) to (AGS) requires *inference*. But inference itself involves following a rule: inference is an example of rule-following *par excellence*. In our (EMAIL) example, the relevant inference involves following MP* as the rule. But

if the (INTENTION VIEW) is correct, then following MP* too requires inference: it requires forming an intention with MP* as its content, and then a belief to the effect that the antecedent of the rule is satisfied. But this lands us, once again, on a vicious infinite regress.

We have seen, moreover, that Miller's solution to the inference problem is the Wittgensteinian one of giving up on the idea that rule-following requires *interpretation*. Once we give up on the idea that rule-following requires interpretation, we need not appeal to a rule over and above (EMAIL) such as MP*: we look to (EMAIL) in a given case and *do something* – respond to the email. It is this solution which might give us the tools to respond to G&W.

Recall that, as G&W stress, any plausible account of rule-following must accommodate a substantive distinction between being guided by a rule, on the one hand, and merely acting in accordance with one, on the other. It is the (INTENTION VIEW) which provides us with this distinction: we follow rules by forming intentions to uphold patterns in our thought or behavior and then acting on those intentions. In following the (EMAIL) rule blindly (where 'blind' contrasts with 'based on interpretation'), the subject follows (EMAIL) and coincidentally *conforms* to MP*: she is not 'really' following MP* as she would be on Boghossian's account.

Recall further that G&W's (2015) exchange with Steglich-Petersen (2013) ends in G&W disputing the idea that there is a *wider* or *more inclusive* notion of guidance on which norms can guide belief. Consider again (D*), the 'causally fortified' version of (D):

> (D*): Acceptance of a norm *N* of the form "In *C*, do *X*" can influence a subject *S*'s behavior with respect to *X* in the sense of making a difference to *S*'s *X*-ing only if *S* can follow *N*.

G&W's objection to (D*) was that even if we read 'influence' as implying some belief-inducing causal mechanism which makes a difference in belief formation,

it is intuitively unclear that this influence is *sufficient* for rule-following, even if it is *necessary*. (This is suggested by the 'only if' formulation in (D*).) G&W, we have seen, take this point to be one more instance of another point familiar from attempts at causal analyses of intentional phenomena. The idea is that, for something to be instance of rule-following, it is not sufficient for acceptance of the rule to causally influence the subject's behavior: rather, acceptance of the rule needs to influence subjects' behavior *in the right way*. But 'the right way', G&W maintain, is imprecise, and so without further argument they cannot grant the normativist (or at least Steglich-Petersen) the idea that there is a wider notion of guidance on which norms can guide.[82]

Miller's solution to the inference problem might help us flesh out 'the right way' in such a way that allows us to block G&W's argument against the idea that there is a wider notion of guidance on which norms can guide. Recall that Miller asks what, if not interpretation of the rule, mediates the 'causal transition' between (INTENTION) and (BELIEF) to (AGS). In our (EMAIL) example, according to Miller (2015: 407), what happens is this: S is in an intentional state (INTENTION) with the content "If you receive an email, answer it!" which combined with (BELIEF) (i.e. "e is an email") causes S to reply to the email *in the right kind of way*. This causal relation is "set up and sustained by facts about custom, practice and training" (Miller 2015: 407). If we can give an account of the way in which acceptance of the rule can causally influence the subject's behavior *in the right kind of way*, then we should be able see how acceptance of the rule is – contra G&W – *sufficient* for rule-following.

---

[82] Note that Steglich-Petersen (2006) has defended the teleological view on which 'belief aims at truth' is interpreted literally. As such, he accepts the truth norm, only denies that it is categorical. I think we can appeal to blind rule-following (where 'blind' contrasts with 'based on interpretation') in order to block G&W's argument against SP-guidance. If the below considerations are correct, then we will have an account of the way in which norms might guide belief, though the question of whether norms can guide belief is quite independent of the question whether they are categorical.

What, then, do we mean by 'the right kind of way'? In fleshing this out, Miller considers a potential objection to his account, according to which the role assigned to the acceptance of a given rule does not provide a satisfactory response to Boghossian's inference problem. Miller (2015: 412) suggests that there are three different ways to understand the role an intention plays when applying a rule in a given case:

(a) It is something to be interpreted; or

(b) It is something that merely causes an action; or

(c) It is something that is not interpreted and something that does not merely cause an action, but rationally results in an action.

The objection, then, is that our story about blind rule-following (where 'blind' contrasts with 'based on interpretation' rather than 'intentional') entitles us only to (b). What we require for a satisfactory response to Boghossian, however, is (c). According to Miller, the question whether we are entitled to (c) or at most (b) "would be urgent for a view on which following a general rule in a given instance consisted in that intention combining with a belief that its triggering condition obtains to cause an action of the sort the rule demands" (2015: 413). For we might get deviant causal chains in following rules. To use our (EMAIL) example, suppose I have (INTENTION) and (BELIEF), but that I get confused or flustered when I have to respond to emails immediately upon receipt, and so start to get upset. Not wanting to be upset, and knowing that I can get rid of the feeling by hitting random keys in the email box and then hitting reply, I do so. Now, Miller points out that in this case, while (INTENTION) and (BELIEF) *cause* me to reply to the email, I do not *follow* the (EMAIL) rule. That is, my action (replying to the email) is caused by (INTENTION) and (BELIEF), but (INTENTION) and (BELIEF) do not *rationalize* my action. How, then, can our story about the causal mediation of (INTENTION) and (BELIEF) to (AGS) entitle us to (c) rather than merely (b)?

In response to this worry, Miller (2015: 413-14) argues that the foregoing story does entitle us to (c). For the possibility of deviant causal chains is precluded by

the presence of 'in the right kind of way': "This requirement excludes cases of deviant causal chains and ensures that the relevant intentional states do not merely cause the relevant action, but rather rationally result in it" (Miller 2015: 413). Miller notes that there are two pertinent questions here:

(i) Is the inclusion of 'in the right kind of way' problematic in the context of our replies to Boghossian?

(ii) How does the inclusion of 'in the right kind of way' square with the emphasis given to the notions of training, custom and practice in the Wittgensteinian alternative to the conception of rule-following as necessarily involving interpretation?

Miller begins with (ii). Consider first the following passage from Wittgenstein's *Philosophical Investigations*:

Let me ask this: what has the expression of a rule – say a signpost – got to do with my actions? What sort of connection obtains here? – Well, this one, for example: I have been trained to react in a particular way to this sign, and now I do so react to it.

But with this you have pointed out only a causal connection; only explained how it has come about that we now go by the signpost; not what this following-the-sign really consists in. Not so; I have further indicated that a person goes by a signpost only in so far as there is an established usage, a custom (1953: §198).

Miller (2015: 413-14) suggests that what Wittgenstein has in mind here is that we earn the right to think of the 'causal transactions' involved in one's encounters with the signpost as genuine instances of rule-following (rather than instances of a mere causal process) when we place an appropriate emphasis on the notions of training and custom. Miller writes:

An intention to uphold a rule combining with a suitable belief to cause an action 'in the right kind of way' is a matter of the relevant agent's having been suitably trained or initiated into an appropriate custom or practice. So the exclusion of deviant causal chains via the inclusion of 'in the right kind of way' in the story above meshes well with an account

122

of rule-following in which interpretation is superseded by training, custom and practice (2015: 414).[83]

This answer to question (ii) gives way to an answer to question (i). We are not attempting to give a *causal analysis* of the notion of intentional action, Miller notes, and so the key question is really this: "Does availing ourselves of a notion of training (or custom or practice) which excludes deviant causal chains in our account of rule-following incur philosophical obligations *in addition to* those that we already have in virtue of subscribing to a non-reductionist view of what constitutes the fact that we are following one rule rather than another?" (2015: 414). The answer to this question, according to Miller, is *negative*, since difficulties non-reductionist conceptions of content face are far from trivial.

Now, how exactly do the foregoing considerations help us to respond to G&W? Recall that Steglich-Petersen contrasts the truth norm with a 'pleasantness norm':

(PLEASANT): S ought to believe that p if and only if believing that p would be pleasant.

Steglich-Petersen's point was that, when subjects accept (TRUTH), rather than (PLEASANT), they end up with different beliefs. Therefore the norms in question can influence S's behavior. Therefore norms can guide belief (SP-guidance). In order to deal with G&W's objections to this wider notion of guidance, we can tell a story similar to our (EMAIL) case. Suppose the subject forms her beliefs by following (TRUTH). First, she will need to form the following intention to comply with the instructions of the norm:

(INTENTION*): For all p, if p is true, believe that p!

---

[83] Miller notes that, given that we are assuming the challenges to a non-reductionist conception of content as a response to KW's sceptic can be met, the notions of training and custom employed here need not be characterized in terms that do not presuppose the notions of meaning and understanding (2015: 414). See also McDowell (1998: 276).

And we combine (INTENTION*) with the following belief:

(BELIEF*): p is true.

The combination of (INTENTION*) and (BELIEF*) yields the action-guiding state, i.e. a state with the content:

(AGS*): Believe that p!

When we give up on the idea that rule-following requires interpretation, we can say that subjects have been *trained* to follow norms like the truth norm in the appropriate way.[84] What mediates the causal transition 'in the right kind of way' from (INTENTION*) and (BELIEF*) to (AGS*) is a matter of the subject's having been initiated into a *custom* or *practice* of believing only the truth. If, by contrast, she were initiated into the custom or practice of only believing what it would be pleasant to believe, then she would end up with wildly different beliefs.[85] It seems to me, then, that Miller's story about the causal transition between the relevant intention and the relevant belief to the action guiding state – when the causal mediation occurs 'in the right kind of way' – allows us to block G&W's argument against a wider or more inclusive notion of guidance. For having fleshed out 'the right kind of way', it may be that acceptance of the rule causally

---

[84] Granted, I don't think objective norms like (TRUTH) guide *directly*. Rather, as hinted at in Chapter 2, I think objective doxastic norms guide *indirectly* via the epistemic norms. Moreover, I think normativism is more plausible when construed in terms of the *knowledge* norm, rather than the truth norm. I only rely on (TRUTH) at present for illustrative purposes. I'll offer an (albeit brief) account of guidance and the knowledge norm in §3.6 below.

[85] It is plausible, however, that there are cases in which some proposition p is both true and pleasant to believe, and cases in which p is true but unpleasant to believe. Suppose, for instance, that you are diagnosed with cancer. Presumably, while it is true to believe that you have cancer, believing that you have cancer would be unpleasant to believe. Your being initiated into a practice of believing only the truth means you *follow* (TRUTH) and do not conform to (PLEASANT). Suppose, on the other hand, that your doctor gives you a clean bill of health. Believing that you are healthy in this case is both true and presumably pleasant to believe. In this case, you follow (TRUTH) because you have been initiated into the practice of believing the truth, and therefore form an intention to comply with the instructions of the rule. However, you also coincidentally *conform* to (PLEASANT) in this case. However, I actually do not think we can follow (TRUTH) in such a direct way, but we are assuming (TRUTH) for the sake of argument. In §3.6 below, I will give a sketch of the way in which the knowledge norm defended previously might guide belief indirectly via the epistemic norms.

influencing the subject's behavior is jointly necessary and sufficient for rule-following. We might, therefore, modify (D*) accordingly (note the biconditional):

(D**): Acceptance of a norm *N* of the form "In *C*, do *X*" can influence a subject *S*'s behavior with respect to *X* in the sense of making a difference to *S*'s *X*-ing *if and only if S* can follow *N*.

G&W maintain that (D*) gets things backwards: "guidance implies influence, but influence can fall short of guidance" (2015: 278). But if our foregoing story about acceptance of a given rule causally influencing the subject's behavior (in the right kind of way) is correct, then as per (D**), influence implies guidance. Just as Miller argued that giving up on the idea that rule-following requires interpretation stops Boghossian's inference problem in its tracks, to give up on the conception of rule-following as interpretation likewise stops in its tracks G&W's argument that there is no wider notion of guidance on which norms can guide belief.

To conclude our response to G&W, then, it is worth noting that G&W reject an appeal to blind rule-following as a response to their arguments against normativism. They write:

> Although the detailed exegesis of Wittgenstein's discussion is much disputed, it is clear that he puts forth a regress argument of some sort, arguing that if one takes rule-following always to involve interpretation then following a rule will be impossible. In response to this problem, it has been suggested that there is a form of rule-following that does not involve any sense of intentionally conforming to the rule: *blind* rule-following. Blind rule-following does not involve any intentional condition, and yet is supposed to be a genuine species of rule-following, distinct from mere brute reactions […] Giving up on the intentional condition on rule-guidance, however, brings us back to square one. There must, we said, be a substantive distinction between mere accordance with a rule, mere regularity, and genuine rule-guidance. How is this distinction to be substantiated once we give up on the most natural and intuitive understanding of it? (2013: 96; cf. 2010b: 160-64).

It is quite clear from this passage, however, that for G&W, 'blind' contrasts with 'intentional'. If we follow rules blindly in this sense, then we lose the distinction

between being guided by a rule and merely acting in accordance with one. But if our foregoing considerations are correct, then we should read 'blind' as contrasting with 'based on interpretation'. Doing so allows us to *preserve* the intentional condition on rule-following, and so it accommodates the sought-after distinction between being guided by a rule and merely acting in accordance with one. An appeal to blind rule-following, then, might allow us to block some of G&W's arguments against normativism, at least when 'blind' contrasts with 'based on interpretation' (which if Miller (2015) and McDowell (1998) are right, is the proper way to read Wittgenstein).[86]

3.5 TOPPINEN ON NORM GUIDANCE

As we saw in the last section, an appeal to blind rule-following might provide us with the tools to block some of G&W's arguments against normativism. In this section, let us look at another potential way of blocking their arguments, due to Teemu Toppinen (2015). Toppinen (2015: 398) reconstructs G&W's no guidance argument as follows (assuming, again, that (TRUTH) is the proper way to formulate the truth norm):

(1) If (TRUTH) is constitutive of belief, then it must be possible for (TRUTH) to guide belief formation.

(2) In order for (TRUTH) to guide belief formation, it must be possible for one to form a belief as to whether p on the basis of (TRUTH).

(3) In order to form a belief as to whether p on the basis of some conditional norm such as (TRUTH), one would first have to form a belief as to whether the 'antecedent condition' of the relevant norm holds.[87]

---

[86] Again, we are assuming that a non-reductionist conception of content in this account of blind rule-following. We have seen that there are non-trivial issues with this, and so the present account is necessarily incomplete. Since I could not hope to deal with these issues properly here, I am only attempting to put the ball in the anti-normativists' court.

[87] Likewise, in order to act on the basis of a norm such as 'Buy low, sell high', one would need to form (among other things) a belief about whether the market is at a low.

(4) So, in order to form a belief as to whether p on the basis of (TRUTH), one would first have to form a belief as to whether the antecedent condition of (TRUTH) holds; that is, as to whether p.

(5) But (TRUTH) cannot offer any guidance for belief formation for someone who has already formed a belief as to whether p.

(6) Therefore, (TRUTH) cannot guide belief formation.

(7) Therefore, (TRUTH) is not constitutive of belief.

As we saw in the foregoing sections, a number of defenders of normativism argue that we do not follow (TRUTH) directly; rather, (TRUTH) guides belief formation *indirectly* via the subjective epistemic norms (e.g. Boghossian 2003, 2005). Call this the *indirect guidance argument*. The idea is that you ought to believe in accordance with your evidence, for example, because doing so is more likely to yield true beliefs, i.e. believing in accordance with evidence is the best means of conforming to norms like (TRUTH). But G&W give the following objection to this sort of strategy:

> Of course, there is good reason to believe that there is a connection between truth (in the actual world) and the 'epistemic norms': normally, following the epistemic norms will lead to truth. For instance, I am more likely to end up with true beliefs if I believe that which I have evidence for, than that which I do not have evidence for. But this connection is purely contingent, whereas what is needed here is a constitutive, or metaphysical, connection such that it can be held that being guided by the objective norm just *is* being guided by the subjective norms (2009: 44-5).

To make matters simple, call this *the contingency objection*. G&W also lodge another, related objection to the indirect guidance argument. They argue that truth norms and epistemic norms imply *two different notions of correctness* (2009: 44). Recall our evidential norm:

(EVIDENCE) S ought to believe only that which is supported by the available evidence.

It is possible for me to believe that which is false on excellent evidential grounds, or that which is true on flimsy evidential grounds. But in both cases, (TRUTH) and (EVIDENCE) give different answers to the question of whether my belief is correct. The normativist wants to show that false beliefs are incorrect. But this conclusion is not supported by (EVIDENCE). For according to (TRUTH), a belief is incorrect when its propositional content is false; whereas according to (EVIDENCE), a belief is incorrect when it is not supported by the evidence. Therefore, according to G&W, the indirect guidance strategy does not support the normativist's view that false beliefs are incorrect, "since the notion of incorrectness implied by [EVIDENCE] does not coincide with that of falsity" (G&W 2009: 44).

Toppinen argues that the contingency objection to the indirect guidance argument is unpersuasive. Suppose that the relevant epistemic norms – such as (EVIDENCE) – are contingent, i.e. they vary from one possible world to another.[88] If we assume this, it is true that being guided by (TRUTH) cannot consist in being guided by a particular epistemic norm like (EVIDENCE). Moreover, the relevant epistemic norms are not on this assumption constitutive of belief. Toppinen writes:

> Yet it could be that to be guided by [TRUTH] just is, roughly, to be disposed to follow whichever epistemic norms turn out to provide effective means of getting at the truth, where this would also involve being disposed to readjust one's habits of belief formation in case one would be doing badly in terms of tracking the truth. The fact (if it is a fact) that the relevant epistemic norms are contingent, varying from one possible world to another, seems simply irrelevant to the question of whether [TRUTH] can guide us through them (2015: 399).

Consider an analogy with objective consequentialism. According to the objective consequentialist, the morally right course of action is the one which *in fact* produces the best consequences.[89] The *objective standard* of right action, then, is a

---

[88] Granted, G&W say here that the *connection* between the objective norm of truth and the subjective epistemic norms is contingent. In more recent work, however, they do claim that the epistemic rules themselves vary from one possible world to another (2013: 88).

[89] See Railton (1984) and Driver (2012: Chapter 5) for discussion.

standard according to which actions are to be evaluated – the morally right ones are the ones which do in fact produce the best results. Objective consequentialism has been criticized on the grounds that agents could not or should not be *guided* by the objective standard. For one thing, we are epistemically limited creatures, and so we are not always in a position to choose the course of action which will in fact produce the best results. Moreover, one might think that being guided by the objective standard is incompatible with being a virtuous or admirable person. For example, you might fail to perform acts of friendship *for the right reasons*; there might be cases in which you perform those actions only because doing so produces the best results.

Objective consequentialists – in response to these criticisms – typically suggest that the objective standard guides *indirectly* via other norms. They distinguish between a *criterion of rightness* and a *decision-making procedure* (as we briefly saw in Chapter 2 with Hattiangadi's defense of (KNOWLEDGE***)). For example, you might decide that the best means of conforming to the objective standard is to maximize expected value; or you might decide to rarely pay attention to moral norms and instead cultivate dispositions to be motivated by, for example, responsiveness to the needs of one's friends (Toppinen 2015: 400). Toppinen writes:

> The objective standard could still have a role to play. Acting on other sorts of grounds could be justified with reference to this standard; it would be possible to occasionally check whether one's motivational structures are acceptable by evaluating them against this standard. Now, it is far from clear, of course, that objective consequentialism has the resources to offer a wholly acceptable account of the action-guiding role of moral theory. But if it is so much as possible to be indirectly guided by the objective consequentialist's standard of rightness, then a similar story should be available also in the case of belief-guidance through [TRUTH] (2015: 400).

In their 2013 paper, G&W frame the contingency objection somewhat differently. On their view, one of the reasons the indirect guidance approach is problematic is that it turns norms like (EVIDENCE) into mere *instrumental* rules:

> [I]f we try to construe the matter in terms of rules, this means that the rules in question not only vary from world to world but are bound to be purely *instrumental* rules. On the assumption that the rules of rationality actually are truth-conducive, the normative force derivable from the aim of truth be would be no different in principle from that of 'imperatives' like 'If you want to go for a ride, you ought to fill up the car'. Since there are instrumental rules contingent on any aims whatsoever, this would trivialize the claim that belief is essentially rule-guided (G&W 2013: 88).[90]

But Toppinen argues that this objection to the indirect guidance argument is also unsuccessful. Suppose that your aim is to always believe what you want to be true. Granted, there are instrumental rules which are contingent on this aim. Toppinen argues, however, that this does not mean that the rules which derive from the aim of believing what you want to be true are on a par with the rules that derive from (TRUTH):

> The normative force derivable from the aim of believing the truth is *very different* from that derivable from the aim of believing what one would want to be true. This is so quite simply because [TRUTH] is true (or so we are supposing), whereas the claim that one always ought to believe what one wants to be true is not. The relevant epistemic rules are contingent and instrumental, alright. But they are not contingent on what aims we happen to have; they derive their normative authority from a true normative principle which is essentially linked to belief. Or that is the idea (Toppinen 2015: 400).

Toppinen again draws an analogy with objective consequentialism. The objective consequentialist says that we ought to perform the action with will in fact maximize the good or produce the best consequences. But suppose that the principle according to which we ought to maximize the good can only guide us indirectly via *contingently* true principles which might require us to maximize expected value or follow the Ten Commandments (for example); but this depends of course on the consequences of following these principles. According to Toppinen, the

---

[90] David Papineau (1999, 2013) endorses a similar view. According to him, there is no distinct species of doxastic normativity; rather, what appear to be doxastic norms just reduce to instrumental norms. On Papineau's view, prescriptions arise when some end is *valuable*, only there is no distinct species of doxastic value that yields doxastic *oughts*. Rather, doxastic *oughts* arise from moral, personal, or even aesthetic value. See Tanney (1999) for a response to Papineau (though not for a defense of normativism).

fact that these principles are contingent "would not trivialize the link between moral rightness and action-guiding rules" (2015: 400). It may be that we ought to maximize expected value, or it may be that we ought to follow the Ten Commandments, depending on the consequences of following the rules in question, "Yet the normative force of these rules would be very different from that of any old hypothetical imperatives, thanks to its deriving from the non-contingent requirement to do the right thing, or to maximize the good, as the objective consequentialist believes" (Toppinen 2015: 400-401).

Thus it looks like one of G&W's main objections to the indirect guidance argument – specifically their contingency objection – is unsuccessful. Toppinen also criticizes the regress of motivations argument and offers his own account of how norms can guide belief. Here is how Toppinen (2015: 402-3) reconstructs the regress of motivations argument:

(i) If certain norms of rationality, R, are essential to belief, then believing that p must be guided by R.

(ii) If believing that p is guided by R, then the R-norms play a role in an *intentional explanation* of believing that p.

(iii) The R-norms play a role in an intentional explanation of believing that p by playing a role in *motivating* believing that p.

(iv) Any motivational explanation, in terms of R, for believing that p will appeal to a prior belief that believing that p is in accordance with R.

(v) The formation of this prior belief, too, must be guided by the rules of rationality, and so an infinite regress ensues. This is bad, and so:

(vi) Norms R are not essential to belief.

Toppinen claims that premises (i) and (ii) of the regress of motivations argument are, on the face of it, highly dubious: "It is not plausible that normativists would be committed to holding that any belief must be guided by the R-norms in the sense of having been *formed* on the basis of R" (2015: 403). However, it is not clear why Toppinen thinks this. Nevertheless, Toppinen suggests what he takes

to be a more plausible way of understanding guidance: in order for S to believe that p under the guidance of R, S's belief that p "must be 'based' on S's acceptance of R in the sense that S *could* arrive at a belief that p, and is *disposed to maintain* this belief, thanks to her accepting R" (2015: 403). Understood this way, Toppinen grants (i) and (ii).

Toppinen suggests rejecting premise (iii). Recall that, according to G&W (2013: 94), on the belief-desire model of intentional explanation, an intentional explanation of believing something requires at least two components: the agent's acceptance of the rule in the 'motivational slot', and a *belief* to the effect that some specific performance is in accordance with the rule. The way in which norms guide belief can then be represented as follows (where the attitude types are in small caps, the contents of the attitudes in the brackets, and the arrow signals a relation of rational explanation):

(P_1) DESIRE (I believe what is in accordance with R)

(P_2) BELIEF (Believing that p is in accordance with R)

$\rightarrow$

(P_3) DESIRE (Believing that p)

$\rightarrow$

(C) BELIEF (p).

But the problem G&W raise arises from (P_1). A desire needs to be 'channeled' through a belief (P_2) in order to have motivational impact, and this leads to a problematic regress. The normativist's idea is that belief formation is rule- or norm-guided, and since (P_2) is a belief, we need to tell a story about how we are motivated to have it. That is, we need:

(C_1): DESIRE (I believe that p is in accordance with R).

And then:

(P₄): BELIEF (Believing that (believing that p is in accordance with R) is in accordance with R).

And so we need:

(C₂): DESIRE (Believing that (P₄)).

But then we need:

(P₅): BELIEF (Believing that (P₄) is in accordance with R).

But again, we need:

(C₃): DESIRE (I believe that (P₅)).

And so on *ad infinitum*. We can never reach (C): an infinite regress ensues. Toppinen suggests, in order to deal with this problem, that rule-acceptance in the present context should be understood as a kind of *inferential habit*:

> [I]nferential habits are intentional states which are in one important respect like desires or wants, and in another respect unlike them. They are like desires in that they, too, seem to be 'pro-attitudes' which have the '*world-to-mind direction of fit*'. Very roughly, whereas a central function of a belief is to fit the facts, a central function of a desire is to produce action which would change the facts so as to be in accordance with its content. Likewise, the function of the *modus ponens* inferential habit is to change the facts – those concerning the belief patterns of the thinker – so as to accord with what is required by the habit. However, unlike desires, habits of inference are not *belief-channeled*. If I have the *modus ponens* habit of inference, I do not need, in order to believe that p on the basis of my believing that q and that if q then p, to believe that by believing that p I believe in accordance with *modus ponens* (Toppinen 2015: 404; cf. Pettit 1993: 18-19).

Toppinen's idea is that, in light of R, thinking that one ought to believe that p can be understood as having some inferential habit functioning to produce the belief that p, given a 'suitable input'. For example, S might believe that p and that p implies q, and so think that, in virtue of having the habit of inferring in accordance with *modus ponens*, she ought to believe that q.

But Toppinen points out that we still need to distinguish being guided by/following a norm and merely acting in accordance with it. If, for example, one thinks that one ought to follow *modus ponens*, this cannot simply consist in being disposed to form beliefs in accordance with *modus ponens*. Rather, as G&W are wont to point out, the *modus ponens* rule would need to *make a difference* in belief formation; that is, the belief must be formed *because* of the rule. Toppinen writes:

> What more is needed [to distinguish conformity to the rule from genuine guidance]? A suggestion roughly along the following lines would seem to hold some promise: following the *modus ponens* rule, rather than simply conforming to it, consists in having a complex dispositional state, which involves not only being disposed to conform to the rule, but also being disposed to 'go Oops!' and to correct one's beliefs upon failing to do so. This picture could be supplemented in various ways. Going Oops! might be taken to involve having certain kinds of feelings as well as directing one's attention in various ways. Also, perhaps further dispositions should be appealed to in order to distinguish between being guided by a rule in the light of normative thought, on one hand, and following a rule without engaging in any specifically normative thinking, on the other (2015: 404).

According to Toppinen, this suggestion would amount to an *expressivist* account of 'epistemic ought thought'. This, Toppinen acknowledges, would be highly controversial. Nevertheless, Toppinen claims that this might be a way in which norms guide belief. Suppose I believe that p on the basis of my thinking that I ought to believe that p, where the judgement that I ought to believe that p is understood as a state of a habit of inference such as *modus ponens*. Now we have the following reasoning:

(P$_1$*) INFERENTIAL HABIT (When in states BELIEF (If q then p) and BELIEF (q), form BELIEF (p)).

(P$_2$*) BELIEF (If q then p).

(P$_3$*) BELIEF (q)

$\rightarrow$

(C*) BELIEF (p)

On this model, believing that p on the basis of having the relevant inferential habit does not require that I have some instrumental belief (e.g. (P_2) above), and so we do not lapse into regress. Thus Toppinen has offered an account of the way in which truth norms might guide belief which does not succumb to the objections of G&W.

However, Toppinen's own positive account of the way in which norms might guide belief is likely to face significant difficulties. We can accept his replies to G&W's objections to normativism, but we will run into trouble if we accept his positive account. For one thing, Toppinen relies on the truth norm. In the previous Chapter, we saw that truth norms face significant difficulties: even if, as Toppinen argues, the truth norm does not succumb to the no guidance argument, it will still succumb to the formulation problem. The upshot of this, recall, was that truth norms are not conceptual truths because *they are not truths*. In the previous Chapter, I argued that the knowledge norm (KNOWLEDGE**) does not succumb to the formulation problem (see also §3.6 below).

Even if Toppinen could somehow dissolve the formulation problem, there is a further difficulty with his account. The difficulty is that, according to him, the way in which norms guide belief amounts to a kind of *expressivism*. As Toppinen says in the passage quoted above (2015: 404), his picture of guidance could be supplemented in various ways. But where exactly does the expressivism come in – in the *picture* of guidance, or the *supplement* to the picture? We might find the expressivism in at least three places:

(i) The relevant doxastic judgement expresses a disposition to conform to the norm; or

(ii) The doxastic judgement expresses a disposition to correct one's belief and 'go Oops!' if the belief doesn't conform to the norm; or

(iii) When 'Oops' expresses some conative feeling or sentiment.

However, regardless of where exactly the expressivism surfaces,[91] Toppinen's positive account of how norms guide belief is likely to face significant difficulties simply in virtue of its expressivism. For expressivism – a form of non-cognitivism – faces a number of well-known difficulties. I will not rehearse these difficulties here, though it is perhaps worth mentioning just one of the more significant ones.

Perhaps the most influential objection to non-cognitivism, and expressivism, is the so-called *Frege-Geach problem*, which Peter Geach (1960, 1965) attributes to Frege's *Begriffsschrift*. The idea is that a non-cognitivist construal of moral language in 'unasserted contexts' (e.g., in the antecedents of conditionals) undermines obviously valid patterns of inference, such as *modus ponens*. Since, according to the non-cognitivist, moral language is not cognitive (i.e. it does not express propositions), we *equivocate* different senses of moral terms in ordinarily valid arguments. For example, according to the non-cognitivist, in the sentence:

(1) Murder is wrong

'Wrong' expresses some conative feeling or sentiment. But when 'wrong' occurs in the following unasserted context:

(2) If murder is wrong, then getting little brother to murder is wrong.

'Wrong' does not express a conative attitude – it is unasserted, embedded within a conditional, and so (2) itself does express a cognitive attitude. If (1) expresses a conative attitude, and (2) a belief, then it does not follow that:

(3) Getting little brother to murder is wrong.

---

[91] In personal correspondence, Toppinen suggests the expressivism in his account surfaces in all three places; that is, in (i), (ii), and (iii).

For we *equivocate* two different senses of 'wrong' in (1) and (2), respectively.[92] Whereas the meaning of (1) is given in terms of the conative mental state expressed by the judgement that murder is wrong, the meaning of (2) is given in terms of its truth-conditions. More, of course, needs to be said, but this is one non-trivial problem for expressivism/non-cognitivism. And others abound.[93] It seems to me, then, that if normativists must appeal to expressivism in order to give a plausible account of the way in which norms guide belief, then so much the worse for normativism.

3.6 INDIRECT GUIDANCE: THE KNOWLEDGE NORM

In the last two sections, we saw that there might be two ways of blocking some of G&W's arguments against normativism. In this section, I want to sketch – albeit briefly – a cognitivist (that is, non-expressivist) account of the way in which norms might guide belief.[94] This account is only provisional, and as such will be subject to a number of provisos.

To begin with, recall that many normativists (such as Boghossian 2005) do not think we follow the truth norm directly; rather, the norm can only be followed by following the subjective epistemic norms, such as norms of evidence or rationality. But as others have pointed out (Hattiangadi 2010; Glüer and Wikforss 2013; Gibbons 2013: Chapter 1), the biggest hurdle this strategy faces is

---

[92] Blackburn (1984) and Gibbard (1990) each have highly controversial responses (respectively) to the Frege-Geach problem. I cannot go into the details of these responses here, though see Miller (2013: Chapters 4 and 5).

[93] See, e.g., Jackson and Pettit (1998) and Miller (2013). As noted in Chapter 1, Jackson (2000) argues that the combination of non-cognitivism and the normativity of belief yields the paradoxical result that there are no beliefs and no believers. If Jackson is right, Toppinen is likely to face this difficulty too, since Toppinen combines expressivism with normativism about belief (though I cannot go into the details of this objection here). See also Hazlett (2013: §9.2.2) for a (sympathetic) discussion of epistemic expressivism.

[94] For reasons of space, this account is necessarily incomplete; I am forced to leave things somewhat vague and speculative. However, leaving things open-ended like this shouldn't be too problematic given the goal of this (and the last) chapter, which is that of *defending* normativism against objections, rather than providing a substantial positive normativist theory. Moreover, I conclude this section with a sketch of the sort of philosophical work that would need to be done in order to make the present account convincing.

the fact that (TRUTH) and its variants can clash with the epistemic norms. Evidence, for example, can be misleading: you might have excellent evidence for a falsehood, and this casts doubt on the idea that the objective norm of truth *grounds* the epistemic norms and supplies their 'rationale'.

In Chapter 2, we saw that the knowledge norm defended does not succumb to this difficulty in the 'hard cases' – coin-toss and lottery cases in particular. That knowledge norm was:

(KNOWLEDGE**) If S considers whether p, then

(i) S ought to (believe that p) if S is in a position to know that p; and

(ii) S ought not to (believe that p) if S is not in a position to know that p.

Unlike its alethic equivalent (TRUTH+), (KNOWLEDGE**) gives the right results in the hard cases. Even if (KNOWLEDGE**) is a constitutive doxastic norm, we might still think of it as an objective standard or criterion of rightness which sets the standard for which epistemic methods to follow. I think it is plausible, based on our foregoing discussion of blind rule-following – and subject to a number of provisos to be outlined below – that (KNOWLEDGE**) guides belief indirectly via the epistemic norms with which, we have seen, it does not clash. The account will be cognitivist, and so will not run into the difficulties that Toppinen's account will.

To see how the objective norm (KNOWLEDGE**) might guide belief indirectly via the epistemic norms, consider our norm of evidence:

(EVIDENCE): S ought to believe only that which is supported by the available evidence.

In general, evidence is truth-conducive, but it is also knowledge-conducive (cf. McHugh 2011).[95] If you have overwhelming evidence that p, then it is likely that p is true, but it is also likely that you are in a position to know that p. For p is likely to be true, and you obviously have evidential justification for believing that p. It is plausible, therefore, that if you were to believe that p, you would know that p.

Now, given our foregoing discussion of blind rule-following (where 'blind' contrasts with 'based on interpretation') we might give the following account. Suppose the subject has been initiated into a practice of only believing what the evidence supports. She might then form an intention to comply with the instructions of the rule:

(INTENTION**): For all p, if the evidence strongly supports p, believe that p!

And we can combine (INTENTION**) with the following belief:

(BELIEF**): The evidence strongly supports p.

The combination of (INTENTION**) and (BELIEF**) brings the subject into an action-guiding state – a state with the content:

(AGS): Believe that p!

There might be a rule-mediated transition between (INTENTION**) and (BELIEF**) which causes the subject 'in the right kind of way' to be in an action-guiding state with "Believe that p!" as its content. This accommodates the distinction between being guided by a rule and merely acting in accordance with one – acceptance of the rule and intention is what gives us that distinction. In

---

[95] According to Timothy Williamson's knowledge-first approach to epistemology, the totality of one's evidence *just is* the totality of one's knowledge (2000: Chapter 9). I do not think, however, that we need to endorse such a radical view in order to give an account of the way in which the knowledge norm might guide indirectly. It is enough to note that, in general, evidence is knowledge-conducive. Moreover, I will leave vague the notion of evidence we are working with, since whatever view of evidence we endorse is likely to be controversial, and I cannot offer a proper discussion here. For discussion of evidence in the present context, as well as further readings, see Olson (2011) and Shah (2011).

this case, you follow (EVIDENCE), and might coincidentally conform to other rules. But it is plausible that following (EVIDENCE) is a means of following (KNOWLEDGE**), since evidence is in general knowledge-conducive, and we have seen (in Chapter 2) that (KNOWLEDGE**) does not clash with the epistemic norms in the hard cases. And if this is right, then (KNOWLEDGE**) guides belief formation and revision indirectly, via the subjective epistemic norms.

I have said that this account would be cognitivist. But what exactly about it is cognitivist? It is cognitivist insofar as doxastic judgements would express propositions; they would not express some conative mental state. Suppose I judge that:

Sarah believes that snow is white.

My judgement expresses the belief or proposition that Sarah believes that snow is white, and the judgement has the truth-value *true* just in case Sarah does believe that snow is white. This, however, is where we need to bring in provisos, and where substantive philosophical work remains to be done.

In fleshing out the present indirect guidance argument, I have helped myself to the notions of meaning, intention, belief, understanding, and so on. I have assumed that these notions are primitive, *sui generis*, and irreducible: that is, I have assumed a non-reductionist conception of content as an answer to KW's sceptic. This, after all, underlies Miller's debate with Boghossian: the latter wanted to argue that, *even assuming* the challenges to non-reductionism can be met, there is a further 'fatal flaw' concerning the intentional view of rule-following. Miller, we have seen, argued that there is no such further problem.

Even so, we are left with the familiar problems concerning the non-reductionist conception of content. The main proviso I want to add to the present account is that it only works (if it works) given a non-reductionist conception of content. But as Boghossian points out, there are still non-trivial difficulties with (ROBUST REALISM) and non-reductionism in general (1989: 48; 2012: 35-6). We need to

explain, for one thing, what room is left for theorizing about meaning and content once we eschew reductionist theories (such as dispositionalism). Moreover, we need an account of the 'casual and explanatory efficacy' of content properties. And finally, we need to explain our first-person knowledge of them.

Furthermore, there are difficulties with (INTENTION VIEW). For one thing, intentionality has been hard to square with *naturalism*: the naturalist insists that intentions must be shown to be naturalistically reducible. Second, Boghossian claims that there seem to be counterexamples to (INTENTION VIEW), though we need not rehearse them here (2012: 35). And finally, we have already seen that (INTENTION VIEW) seems to require us to reject (MEANING ASSUMPTION), but it is not clear that we can reject it. After all, KW regards (MEANING ASSUMPTION) as non-optional.

In order to complete our present account of guidance, therefore, one would need to show how these difficulties can be met. I think it is plausible that (KNOWLEDGE**) guides belief formation indirectly in the way sketched above, but this will not be satisfactory until the difficulties with non-reductionism that underlie the idea of blind rule-following we are working with are indeed met.

3.7 CONCLUSION

This chapter has considered two potential ways of blocking some of G&W's arguments against normativism. First, we might appeal to blind rule-following (where 'blind' contrasts with 'based on interpretation' rather than 'intentional') in order to block G&W's arguments against the idea that there is a wider notion of guidance on which norms can guide belief. Since this account allows us to preserve (INTENTION VIEW), we can accommodate the distinction between following a rule and merely acting in accordance with one. Second, as Toppinen argued, the fact that the connection between the objective doxastic norm and the subjective epistemic ones is contingent does not mean the former is not capable of guiding us through the latter. However, Toppinen's own account is likely to face difficulties. This chapter concluded with a sketch of a cognitivist account of

the way in which (KNOWLEDGE**) might guide belief indirectly, though signifi-
cant philosophical work remains to be done in order to make that account con-
vincing.

CONCLUSION

According to the moral error theory, there are no moral facts: moral judgements are systematically and uniformly false. Moral realism, by contrast, is the theory that moral judgements purport to state facts, and that some of them are facts: some moral judgements have true contents. This thesis has deployed a companions in guilt argument in order to defend moral realism against Mackie's argument for the error theory: the argument from queerness.

Chapter 1 presented the threat to the error theory. After giving an overview of the argument from queerness and contrasting the error theory with moral realism and ethical non-cognitivism, I appealed to doxastic normativism in order to flesh out the threat. That threat was that, if normativism is true, then the argument from queerness can be run in the case of belief. If a doxastic analogue of the argument from queerness – based on normativism – were sound, then it would support the implausible error theoretic conclusion that doxastic judgements are systematically and uniformly false. The upshot of this is the paradoxical one that there are no beliefs and no believers. To commit to this conclusion would be to commit cognitive suicide; and we saw that a defender of the moral error theory cannot appeal to hypothetical doxastic imperatives in order to assuage the consequences of our doxastic argument from queerness.

Whether the arguments of Chapter 1 succeed, however, depends in large part on whether normativism is true. In Chapters 2 and 3, therefore, I defended normativism against some influential objections to it. Following clarification of what normativism amounts to, Chapter 2 looked at the objections Bykvist and Hattiangadi raise against truth norms for belief. I argued that a particular iteration of the knowledge norm does not succumb to these objections, that it is a plausible candidate for a constitutive norm which serves to distinguish belief from other attitudes, and that it explains why the epistemic norms hold, i.e. it

explains why we ought to seek evidential justification for our beliefs, why our belief formation ought to be rational, and so on. Moreover, the objective knowledge norm defended does not clash with the epistemic norms in the 'hard cases' – in particular, coin-toss and lottery cases.

Chapter 3 examined the objections Glüer and Wikforss raise against normativism. I considered two ways we might block some of their arguments against normativism. The first appealed to blind rule-following (where 'blind' contrasts with 'based on interpretation'). If these considerations are correct, then it is not the case that Glüer and Wikforss have succeeded in establishing that there is no wider or more inclusive conception of rule- or norm-guidance on which norms can guide belief. The second comes from Teemu Toppinen. If Toppinen is right, then Glüer and Wikforss do not succeed in refuting the idea that norms can guide belief indirectly. However, Toppinen's own positive account of the way in which norms might guide belief is likely to face difficulties – in particular, it amounts to a problematic *expressivist* treatment of 'epistemic ought thought'. I concluded with a brief sketch of the way in which the knowledge norm defended in Chapter 2 might guide belief formation indirectly, via the epistemic norms. This account, however, is only provisional, and much philosophical work remains to be done in order to make it convincing.

BIBLIOGRAPHY

Anscombe, G. E. M. 1957. *Intention*. Oxford: Blackwell.

Ayer, A. J. 1946 [1990]. *Language, Truth and Logic*. London: Penguin. 2nd ed.

Baker, Lynn Rudder. 1998. "Cognitive Suicide." In *Contents of Thought: Arizona Colloquium in Cognition*, edited by Robert Grimm and Daniel Merill. Tuscon: University of Arizona Press, 1-30.

Bedke, Matthew. 2010. "Might All Normativity Be Queer?" *Australasian Journal of Philosophy* 88 (1): 41–58.

Blackburn, Simon. 1984. *Spreading the Word: Groundings in the Philosophy of Language*. Oxford: Oxford University Press.

Boghossian, Paul. 1989 [2008]. "The Rule-Following Considerations." Reprinted in *Content and Justification: Philosophical Papers*. Oxford: Clarendon Press, 9–50, to which page references apply.

Boghossian, Paul. 2003. "The Normativity of Content." *Philosophical Issues* 13 (1): 31–45.

Boghossian, Paul. 2005 [2008]. "Is Meaning Normative?" Reprinted in *Content and Justification: Philosophical Papers*. Oxford: Clarendon Press, 95–107, to which page references apply.

Boghossian, Paul. 2008. "Epistemic Rules." Reprinted in *Content and Justification: Philosophical Papers*. Oxford: Clarendon Press, 109–34, to which page references apply.

Boghossian, Paul. 2012. "Blind Rule-Following." In *Mind, Meaning, and Knowledge: Essays on Themes from the Philosophy of Crispin Wright*, edited by Annalisa Coliva, Oxford: Clarendon Press, 27–48.

Brandom, Robert. 1994. *Making It Explicit: Reasoning, Representing and Discursive Commitment*. Cambridge: Harvard University Press.

Brink, David O. 1986. "Externalist Moral Realism." *The Southern Journal of Philosophy* 24 (1): 23–42.

Broome, John. 2000. "Normative Requirements." In *Normativity*, edited by Johnathan Dancy. Malden: Blackwell, 78–99.

Broome, John. 2004. "Reasons." In *Reason and Value: Themes from the Moral Philosophy of Joseph Raz*, edited by R. Jay Wallace, Philip Pettit, Samuel Scheffler, and Michael Smith. Oxford: Clarendon Press, 28–55.

Bykvist, Krister. 2007. "Violations of Normative Invariance: Some Thoughts on Shifty Oughts." *Theoria* 73 (2): 98–120.

Bykvist, Krister, and Hattiangadi, Anandi. 2007. "Does Thought Imply Ought?" *Analysis* 67 (4): 277–85.

Bykvist, Krister, and Hattiangadi, Anandi. 2013. "Belief, Truth and Blindspots." In *The Aim of Belief*, edited by Timothy Chan. Oxford: Oxford University Press, 100–122.

Chrisman, Matthew. 2008. "Ought to Believe." *The Journal of Philosophy* 105 (7): 346–70.

Clark-Doane, Justin. 2012. "Morality and Mathematics: The Evolutionary Challenge." *Ethics* 122 (2): 313–40.

Clipsham, Patrick. 201X. "Is There a Distinct Metaphilosophical Companions in Guilt Argument (And Does It Work)?" Forthcoming in *Philosophia.*

Cowie, Christopher. 2014. "Why Companions in Guilt Arguments Won't Work." *The Philosophical Quarterly* 64 (256): 407–22.

Cowie, Christopher. 2016. "Good News for Moral Error Theorists: A Master Argument against Companions in Guilt Strategies." *Australasian Journal of Philosophy* 94 (1): 115–30.

Cuneo, Terence. 2007. *The Normative Web: An Argument for Moral Realism.* Oxford: Oxford University Press.

Darwall, Stephen, Gibbard, Allan, and Railton, Peter. 1992. "Toward *Fin de Siècle* Ethics: Some Trends." *The Philosophical Review* 101 (1): 115–89.

Das, Ramon. 2016. "Why Companions in Guilt Arguments Still Work: Reply to Cowie." *The Philosophical Quarterly* 66 (262): 152–60.

Das, Ramon. 2017. "Bad News for Moral Error Theorists: There is no Master Argument against Companions in Guilt Strategies." *Australasian Journal of Philosophy* 95 (1): 58–69.

Douven, Igor. 2008. "The Lottery Paradox and Our Epistemic Goal." *Pacific Philosophical Quarterly* 89 (2): 204–25.

Driver, Julia. 2012. *Consequentialism.* Abingdon: Routledge.

Engel, Pascal. 2007. "Belief and Normativity." *Disputatio* 2 (23): 1–25.

Engel, Pascal. 2013a. "In Defense of Normativism about the Aim of Belief." In *The Aim of Belief*, edited by Timothy Chan. Oxford: Oxford University Press, 31–62.

Engel, Pascal. 2013b. "Doxastic Correctness." *Proceedings of the Aristotelian Society, Supplementary Volume* 87 (1): 199–216.

Evans, Matthew, and Shah, Nishi. 2012. "Mental Agency and Metaethics." In *Oxford Studies in Metaethics*, vol. 7, edited by Russ Shafer-Landau. Oxford: Oxford University Press, 80–109.

Finlay, Stephen. 2008. "The Error in the Error Theory." *Australasian Journal of Philosophy* 86 (3): 347–69.

Finlay, Stephen. 2010. "Recent Work on Normativity." *Analysis* 70 (2): 331–46.

Fletcher, Guy. 2017. "Pain for the Moral Error Theory? A New Companions-in-Guilt Argument." *Australasian Journal of Philosophy* 93 (3): 474-82.

Foot, Philippa. 1972. "Morality as a System of Hypothetical Imperatives." *The Philosophical Review* 81 (3): 305–16.

Garner, Richard. 1990. "On the Genuine Queerness of Moral Properties and Facts." *Australasian Journal of Philosophy* 68 (2): 137–46.

Garner, Richard. 2007. "Abolishing Morality." *Ethical Theory and Moral Practice* 10 (5): 499–513.

Geach, P. T. 1960. "Ascriptivism." *The Philosophical Review* 69 (2): 221–25.

Geach, P. T. 1965. "Assertion." *The Philosophical Review* 74 (4): 449–65.

Gibbard, Allan. 1990. *Wise Choices, Apt Feelings: A Theory of Normative Judgement.* Oxford: Clarendon Press.

Gibbard, Allan. 2003. "Thoughts and Norms." *Philosophical Issues* 13 (1): 83–98.

Gibbard, Allan. 2005. "Truth and Correct Belief." *Philosophical Issues* 15 (1): 338–50.

Gibbard, Allan. 2012. *Meaning and Normativity*. Oxford: Oxford University Press.

Gibbons, John. 2013. *The Norm of Belief*. Oxford: Oxford University Press.

Glüer, Kathrin. 1999. *Sprache und Regeln: Zur Normativität von Bedeutung*. Berlin: De Gruyter.

Glüer, Kathrin, and Pagin, Peter. 1999. "Rules of Meaning and Practical Reasoning." *Synthese* 117 (1): 207–27.

Glüer, Kathrin, and Wikforss, Åsa. 2009. "Against Content Normativity." *Mind* 118 (469): 31–70.

Glüer, Kathrin, and Wikforss, Åsa. 2010a. "The Truth Norm and Guidance: A Reply to Steglich-Petersen." *Mind* 119 (475): 757–61.

Glüer, Kathrin, and Wikforss, Åsa. 2010b. "Es braucht die Regel nicht: Wittgenstein on Rules and Meaning." In *The Later Wittgenstein on Language*, edited by Daniel Whiting. Basingstoke: Palgrave Macmillan, 148–66.

Glüer, Kathrin, and Wikforss, Åsa. 2013. "Against Belief Normativity." In *The Aim of Belief*, edited by Timothy Chan. Oxford: Oxford University Press, 80-99.

Glüer, Kathrin, and Wikforss, Åsa. 2015. "Still No Guidance: Reply to Steglich-Petersen." *Theoria* 81 (1): 272–79.

Greco, Daniel. 2015. "Epistemological Open Questions." *Australasian Journal of Philosophy* 93 (3): 509–23.

Greenberg, Alexander. 201X. "Should I Believe All the Truths?" Forthcoming in *Synthese*.

Hattiangadi, Anandi. 2007. *Oughts and Thoughts: Rule-Following and the Normativity of Content*. Oxford: Clarendon Press.

Hattiangadi, Anandi. 2010. "The Love of Truth." *Studies in History and Philosophy of Science* 41 (1): 422–32.

Hazlett, Allan. 2013. *A Luxury of the Understanding: On the Value of True Belief*. Oxford: Oxford University Press.

Heathwood, Chris. 2009. "Moral and Epistemic Open-Question Arguments." *Philosophical Books* 50 (2): 83–98.

Hinckfuss, Ian. 1987. *The Moral Society: Its Structure and Effects*. Canberra: Australian National University.

Hlobil, Ulf. 2015. "Anti-Normativism Evaluated." *International Journal of Philosophical Studies* 23 (3): 376–95.

Horgan, Terry, and Timmons, Mark. 2000. "Nondescriptivist Cognitivism: Framework for a New Metaethic." *Philosophical Papers* 29 (1): 121–53.

Hume, David. 1738 [1978]. *A Treatise of Human Nature*. Edited by L. A. Selby-Bigge and P. H. Nidditch. Oxford: Clarendon Press.

Hursthouse, Rosalind. 2000. *On Virtue Ethics*. Oxford: Oxford University Press.

Husi, Stan. 2013. "Why Reasons Skepticism Is Not Self-Defeating." *European Journal of Philosophy* 21 (3): 424–49.

Ichikawa, Johnathan Jenkins, and Jarvis, Benjamin W. 2013. *The Rules of Thought*. Oxford: Oxford University Press.

Jackson, Frank. 2000. "Non-Cognitivism, Normativity, Belief." In *Normativity*, edited by Johnathan Dancy. Malden: Blackwell, 100–115.

Jackson, Frank, and Pettit, Philip. 1998. "A Problem for Expressivism." *Analysis* 58 (4): 239–51.

Joyce, Richard. 2001. *The Myth of Morality*. Cambridge: Cambridge University Press.

Joyce, Richard. 2011. "The Error in 'The Error in the Error Theory.'" *Australasian Journal of Philosophy* 89 (3): 519–34.

Kalantari, Seyed Ali, and Luntley, Michael. 2013. "On the Logic of Aiming at Truth." *Analysis* 73 (3): 419–22.

Kalderon, Mark. 2005. *Moral Fictionalism*. Oxford: Clarendon Press.

Kant, Immanuel. 1781 [1998]. *Critique of Pure Reason*. Edited and translated by Paul Guyer and Allen W. Wood. Cambridge: Cambridge University Press.

Kant, Immanuel. 1785 [2011]. *Groundwork of the Metaphysics of Morals*. Edited by Jens Timmermann. Translated by Mary J. Gregor. Cambridge: Cambridge University Press.

Keller, Simon. 2017. "Love and the Moral Error Theory: Is Love a Mistake?" *Philosophy and Phenomenological Research* 95 (3): 709–21.

Kelly, Thomas. 2003. "Epistemic Rationality as Instrumental Rationality: A Critique." *Philosophy and Phenomenological Research* 66 (3): 612–40.

Korsgaard, Christine. 1996. *The Sources of Normativity*. Cambridge: Harvard University Press.

Kripke, Saul. 1982. *Wittgenstein on Rules and Private Language*. Oxford: Blackwell.

Lillehammer, Hallvard. 2007. *Companions in Guilt: Arguments for Ethical Objectivity*. London: Palgrave Macmillan.

Littlejohn, Clayton. 2013. "The Russellian Retreat." *Proceedings of the Aristotelian Society, Supplementary Volume* 113 (3): 293–320.

Mackie, J. L. 1977 [1990]. *Ethics: Inventing Right and Wrong*. New York: Penguin.

McDowell, John. 1994. *Mind and World*. Cambridge: Harvard University Press.

McDowell, John. 1998. *Mind, Value and Reality*. Cambridge: Harvard University Press.

McGinn, Colin. 1984. *Wittgenstein on Meaning*. Oxford: Blackwell.

McHugh, Conor. 2011. "What Do We Aim at When We Believe?" *Dialectica* 65 (3): 369–92.

McHugh, Conor. 2012a. "Epistemic Deontology and Voluntariness." *Erkenntnis* 77 (1): 65–94.

McHugh, Conor. 2012b. "The Truth Norm of Belief." *Pacific Philosophical Quarterly* 32 (1): 8–30.

McHugh, Conor, and Whiting, Daniel. 2014. "Recent Work on the Normativity of Belief." *Analysis* 74 (4): 698–713.

Millar, Alan. 2004. *Understanding People: Normativity and Rationalizing Explanation*. Oxford: Oxford University Press.

Millar, Alan. 2009. "How Reasons for Action Differ from Reasons for Belief." In *Spheres of Reason: New Essays in the Philosophy of Normativity*, edited by Simon Robertson. Oxford: Oxford University Press, 140–63.

Miller, Alexander. 2008. "Thoughts, Oughts and the Conceptual Primacy of Belief." *Analysis* 68 (3): 234–38.

Miller, Alexander. 2010. "The Argument from Queerness and the Normativity of Meaning." In *Truth, Existence and Realism*, edited by M. Grajner and A. Rami. Frankfurt am Main: Ontos.

Miller, Alexander. 2013. *Contemporary Metaethics: An Introduction*. Cambridge: Polity Press. 2nd ed.

Miller, Alexander. 2015. "Blind Rule-Following and the 'Antinomy of Pure Reason.'" *The Philosophical Quarterly* 65 (260): 396–416.

Mitova, Veli. 2011. "Epistemic Motivation: Towards a Metaethics of Belief." In *Reasons for Belief*, edited by Andrew Reisner and Asbjørn Steglich-Petersen. Cambridge: Cambridge University Press, 54–74.

Mizrahi, Moti. 2012. "Does 'Ought' Imply 'Can' from an Epistemic Point of View?" *Philosophia* 40 (4): 829–40.

Moore, G. E. 1903 [1993]. *Principia Ethica*. Cambridge: Cambridge University Press. Revised Edition.

Olson, Jonas. 2011. "Error Theory and Reasons for Belief." In *Reasons for Belief*, edited by Andrew Reisner and Asbjørn Steglich-Petersen. Cambridge: Cambridge University Press, 75-93.

Olson, Jonas. 2014. *Moral Error Theory: History, Critique, Defense*. Oxford: Oxford University Press.

Owens, David. 2000. *Reason Without Freedom: The Problem of Epistemic Normativity*. London: Routledge.

Papineau, David. 1999. "Normativity and Judgement." *Proceedings of the Aristotelian Society, Supplementary Volume* 73 (1): 17–43.

Papineau, David. 2013. "There Are No Norms of Belief." In *The Aim of Belief*, edited by Timothy Chan. Oxford: Oxford University Press, 64–79.

Petersen, Steve. 2013. "Utilitarian Epistemology." *Synthese* 190 (1): 1173–84.

Pettit, Philip. 1993. *The Common Mind: An Essay on Psychology, Society and Politics*. Oxford: Oxford University Press.

Pigden, Charles R. 2007. "Nihilism, Nietzsche and the Doppelgänger Problem." *Ethical Theory and Moral Practice* 10 (5): 441–56.

Railton, Peter. 1984. "Alienation, Consequentialism, and the Demands of Morality." *Philosophy and Public Affairs* 13 (2): 134–71.

Railton, Peter. 1986. "Moral Realism." *The Philosophical Review* 95 (2): 163–207.

Railton, Peter. 2000. "Normative Force and Normative Freedom: Hume and Kant, but not Hume *versus* Kant." In *Normativity*, edited by Johnathan Dancy. Malden: Blackwell, 1-33.

Raleigh, Thomas. 2013. "Belief Norms and Blindspots." *The Southern Journal of Philosophy* 51 (2): 243–69.

Raz, Joseph. 1975 [1999]. *Practical Reasoning and Norms*. Oxford: Oxford University Press.

Reisner, Andrew. 2013. "Leaps of Knowledge." In *The Aim of Belief*, edited by Timothy Chan. Oxford: Oxford University Press, 167–83.

Rowland, Richard. 2013. "Moral Error Theory and the Argument from Epistemic Reasons." *Journal of Ethics and Social Philosophy* 7 (1): 1–24.

Rowland, Richard. 2016. "Rescuing Companions in Guilt Arguments." *The Philosophical Quarterly* 66 (262): 161–71.

Sayre-McCord, Geoffrey. 1986. "The Many Moral Realisms." *The Southern Journal of Philosophy* 24 (5): 1–22.

Sellars, Wilfred. 1956. "Empiricism and the Philosophy of Mind." In *Science, Perception and Reality*. Atascadero: Ridgeview Publishing, 127-96.

Shafer-Landau, Russ. 2003. *Moral Realism: A Defense*. Oxford: Clarendon Press.

Shah, Nishi. 2003. "How Truth Governs Belief." *The Philosophical Review* 112 (4): 447–82.

Shah, Nishi. 2011. "Can Reasons for Belief Be Debunked?" In *Reasons for Belief*, edited by Andrew Reisner and Asbjørn Steglich-Petersen. Cambridge: Cambridge University Press, 94–107.

Shah, Nishi, and Velleman, J. David. 2005. "Doxastic Deliberation." *The Philosophical Review* 114 (4): 497–534.

Smith, Michael. 1994. *The Moral Problem*. Malden: Blackwell.

Smithies, Declan. 2012. "The Normative Role of Knowledge." *Noûs* 46 (2): 265–88.

Sorensen, Roy. 1988. *Blindspots*. Oxford: Clarendon Press.

Steglich-Petersen, Asbjørn. 2006. "No Norm Needed: On the Aim of Belief." *The Philosophical Quarterly* 56 (225): 499–516.

Steglich-Petersen, Asbjørn. 2010. "The Truth Norm and Guidance: A Reply to Glüer and Wikforss." *Mind* 119 (475): 749–55.

Steglich-Petersen, Asbjørn. 2013. "The No Guidance Argument." *Theoria* 79 (1): 279–83.

Stevenson, C. L. 1944. *Ethics and Language*. New Haven: Yale University Press.

Stratton-Lake, Philip. 2002. "Introduction." In *The Right and the Good*, by W. D. Ross. Oxford: Clarendon Press, ix–lviii.

Streumer, Bart. 2017. *Unbelievable Errors: An Error Theory about All Normative Judgements*. Oxford: Oxford University Press.

Sturgeon, Nicholas. 1985. "Moral Explanations." In *Morality, Reason and Truth*. London: Rowman and Allanhead, 49–78.

Sutton, Johnathan. 2007. *Without Justification*. Cambridge: MIT Press.

Tanney, Julia. 1999. "Normativity and Thought." *Proceedings of the Aristotelian Society, Supplementary Volume* 73 (1): 45–61.

Toppinen, Teemu. 2015. "How Norms (Might) Guide Belief." *International Journal of Philosophical Studies* 23 (3): 396–409.

Velleman, J. David. 2000. "On the Aim of Belief." In *The Possibility of Practical Reason*. Oxford: Clarendon Press, 244–82.

Wedgwood, Ralph. 2002. "The Aim of Belief." *Philosophical Perspectives* 16: 267–97.

Wedgwood, Ralph. 2007. *The Nature of Normativity*. Oxford: Clarendon Press.

Wedgwood, Ralph. 2013a. "The Right Thing to Believe." In *The Aim of Belief*, edited by Timothy Chan. Oxford: Oxford University Press, 123–39.

Wedgwood, Ralph. 2013b. "Doxastic Correctness." *Proceedings of the Aristotelian Society, Supplementary Volume* 87 (1): 217–34.

Whiting, Daniel. 2010. "Should I Believe the Truth?" *Dialectica* 64 (2): 213–24.

Whiting, Daniel. 2012. "Does Belief Aim (Only) at the Truth?" *Pacific Philosophical Quarterly* 93 (1): 279–300.

Whiting, Daniel. 2013. "Truth: The Aim and Norm of Belief." *Teorema* 32 (3): 121–35.

Wikforss, Åsa. 2001. "Semantic Normativity." *Philosophical Studies* 102 (1): 203–26.

Wikforss, Åsa. 2018. "Does Semantics Need Normativity? Comments on Alan Gibbard's *Meaning and Normativity*." *Inquiry* 61 (7): 755–66.

Williams, Bernard. 1973. "Deciding to Believe." In *Problems of the Self*. Cambridge: Cambridge University Press, 136–50.

Williamson, Timothy. 2000. *Knowledge and Its Limits*. Oxford: Oxford University Press.

Wittgenstein, Ludwig. 1953 [2009]. *Philosophical Investigations*. Edited and translated by G. E. M. Anscombe, P. M S. Hacker, and Joachim Schulte. Oxford: Blackwell. 4th ed.

Worsnip, Alex. 2016. "Moral Reasons, Epistemic Reasons, and Rationality." *The Philosophical Quarterly* 66 (263): 341–61.

Wright, Crispin. 2001. *Rails to Infinity: Essays on Themes from Wittgenstein's Philosophical Investigations*. Cambridge: Harvard University Press.

Wright, Crispin. 2007. "Rule-Following Without Reasons: Wittgenstein's Quietism and the Constitutive Question." *Ratio* 20 (4): 481–502.

Wright, Crispin. 2012. "Replies Part I: The Rule-Following Considerations and the Normativity of Meaning." In *Mind, Meaning, and Knowledge: Essays on Themes from the Philosophy of Crispin Wright*, edited by Annalisa Coliva. Oxford: Oxford University Press, 379–401.

Zagzebski, Linda. 2003. "Intellectual Motivation and the Good of Truth." In *Intellectual Virtue: Perspectives from Ethics and Epistemology*, edited by Michael DePaul and Linda Zagzebski. Oxford: Clarendon Press, 135–54.

Zalabardo, José L. 2010. "Why Believe the Truth? Shah and Velleman on the Aim of Belief." *Philosophical Explorations* 13 (1): 1–21.