



# Durham E-Theses

---

## *Naturalism and the Problem of Normativity*

TURP, MICHAEL-JOHN,DAVID

### How to cite:

---

TURP, MICHAEL-JOHN,DAVID (2011) *Naturalism and the Problem of Normativity*, Durham theses, Durham University. Available at Durham E-Theses Online: <http://etheses.dur.ac.uk/3469/>

### Use policy

---

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in Durham E-Theses
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full Durham E-Theses policy](#) for further details.

**Naturalism and the Problem of Normativity**  
**Michael-John Turp**

**Abstract**

This dissertation explores the way in which normative facts create a problem for naturalist approaches to philosophy. How can lumpy scientific matter give rise to technicolour normativity? How can normative facts show up in the world described from a scientific perspective? In this context, I start by analysing Hume's discussion of 'is' and 'ought', Moore's open question argument, and Kripke's interpretation of Wittgenstein's rule-following considerations. I then look at the nature of philosophical naturalism in detail, arguing that is fundamentally an epistemological commitment to the norms governing scientific publications. I consider the particular examples of Penelope Maddy's approach to naturalising logic and the instrumentalist accounts of epistemic normativity favoured by advocates of naturalised epistemology. I argue, however, that these approaches to naturalising normativity are unsuccessful. In the second half of the dissertation, I develop a novel account of the nature of normative facts and explain how this relates to and resolves some of the difficulties raised in the first half. The account I defend has Kantian foundations and an Aristotelian superstructure. I associate the right with the necessary preconditions for engaging in valuable activity and the good with the satisfaction of the constitutive ends of activities and practices. I explain how my theory can account for epistemic normativity and defend a virtue-based theory of epistemic evaluation. Finally, I argue against desire-based accounts of reasons and in favour of a role for the emotions in normative cognition. The view I defend is intended to be compatible with our best scientific theories. However, it is not naturalistic insofar as it is justified by distinctively philosophical methods and relies on extra-scientific considerations.

**Naturalism and the Problem of Normativity**

**Michael-John David Turp**

**Submitted for the degree of Ph.D. in Philosophy**

**Department of Philosophy  
University of Durham**

**Submitted October 2011**

## CONTENTS

<i>Acknowledgements</i>	v
<i>Introduction</i>	vii
Chapter 1: Three Case Studies in Naturalism and Normativity	1
1.1 Hume on ‘is’ and ‘ought’	1
1.2 Moore’s Open Question Argument	13
1.3 The Rule-Following Considerations	20
1.4 Naturalism and Classical Realism	25
Chapter 2: Naturalism as an Epistemological Stance	33
2.1 Epistemological and Ontological Naturalism	33
2.2 Is Epistemological Naturalism Self-Defeating?	41
Chapter 3: Naturalism and Scientific Practice	47
3.1 The Properties and Methods of Science	47
3.2 Articles as Proxy Indicators of Scientific Practice	51
3.3 Scientific Articles and Objectification	57
Chapter 4: Naturalised Logic and Normativity	64
4.1 Maddy’s Naturalised Philosophy of Logic	64
4.2 Justificatory and Explanatory Reasons	70
4.3 Bridging the Gap with Psychologism	75
Chapter 5: Naturalised Epistemology and Instrumentalism	83
5.1 Naturalised Epistemology and Normativity	83
5.2 Instrumentalist Approaches to Epistemic Normativity	88
5.3 Categorical Epistemic Norms and Trivial Knowledge	97
5.4 Instrumentalism, Contingency and Truth	104

Chapter 6: A Sketch of the Good and the Right	113
6.1 The Logical Grammar of Evaluative Terms	113
6.2 Functions and Natural Goodness	121
6.3 Goodness and the Practice of Chess	126
6.4 Rightness and the Practice of Chess	133
6.5 Non-Contingent Needs, Flourishing and Meaning	138
6.6 Life's Meaning and Value	144
Chapter 7: Truth, Virtue and Epistemic Normativity	151
7.1 True Beliefs and Epistemic Virtue	152
7.2 Demanding the Impossible: Obligation and Evaluation	160
7.3 Knowledge and Practical Interests	163
Chapter 8: Normative Force, Reasons and Desires	168
8.1 Desire-Based Reasons	168
8.2 Desires and the Rationality of Ends	179
Chapter 9: Normativity, Queerness and the Emotions	189
9.1 Mackie's Argument from Queerness	189
9.2 Moral Cognition and the Brain	194
9.3 Emotion and Feeling	196
9.4 Emotional Experience and Intentionality	200
9.5 Emotion and Perception	206
9.6 Conclusion	209
<i>Bibliography</i>	211

*The copyright of this thesis rests with the author. No quotation from it should be published without the prior written consent and information derived from it should be acknowledged.*

## **Acknowledgments**

Preparing a doctoral dissertation is an unusual and challenging task. It would have been an impossible one without the help and support of staff and fellow students at Durham University. I owe particular thanks to my supervisors Jonathan Lowe and Geoffrey Scarre for their intellectual guidance, help and support through a long process. I would also like to offer my thanks to Tony Booth, Matthew Conduct, Owen Earnshaw, Marcos Galounis, Beth Hannon, Simon James, Ian Kidd, Liz McKinnell, Phil Meadows, Donnchadh O’Conaill, Andreas Pantazatos, Duncan Proctor, Soran Reader, Richard Stopford, Tuomas Tahko, Christopher Vaughan and Paul Winstanley. I have benefitted greatly from discussion with all of them. Anything that is right in this dissertation I take full credit for. Any errors, I blame on corrupting influences.

On a more personal level, I am first and foremost very grateful to my wife Ola, whose fortitude in the face of daily philosophical interrogations has been above and beyond the call of duty. I should also like to thank her parents, Ewa and Mariusz, for their friendship, moral and practical support.

I would have liked to have been able to thank my grandmother Dorothy Rachel Turp for her practical wisdom and unflagging sense of humour. Sadly, however, she passed away as I was in the process of writing up. I imagine that she would have been happy that I have finally completed my thesis, but perhaps not overly impressed. “Doctor!” she was fond of saying, “I’ve been D.R. Turp all my life!”

I also thank audiences at the universities of Durham, Leeds, Bucharest and Salzburg for comments and feedback on various sections of the dissertation.

Finally I am grateful for the financial support of the Arts and Humanities Research Council, without which I could not have undertaken this research.

Dedicated to the memory of my father  
Michael Anthony Turp (1951 – 2001)



## **Introduction**

In this dissertation, I explore the sense in which normative facts pose a problem for naturalism. How can lumpy scientific matter give rise to technicolour normativity? How can normative facts show up in the world described from a scientific perspective? These are vague worries and the first half of the dissertation attempts to make it clearer what they amount to. In the second half, I argue for a particular account of the nature of normative facts and explain how this relates to and resolves some of the difficulties raised in the first half. In successive chapters, I sketch an account of the metaphysics, epistemology and motivational force of normative properties, with a particular focus on epistemic normativity. This is an ambitious project and, doubtless, the position I advocate is not right on every point. Nevertheless, working through a plausible, constructive account of normative properties should throw light on the nature of normativity and help to make the problems and prospects of naturalism clearer.

There is of course a long history of philosophers worrying about how to locate the normative in the scientifically given natural world. In chapter one, I look at three influential discussions, namely Hume's discussion of 'is' and 'ought', Moore's open question argument, and Kripke's interpretation of Wittgenstein's rule-following considerations. As well as advancing a partly novel view of Hume's argument, I uncover structural similarities between the three sets of considerations and use them to refine my initial concerns. The question becomes one of how naturalists might identify truth-makers for normative propositions. As these truth-makers have the special property of possessing normative force or authority, I propose the label of value-makers. I restrict my discussion to realist, cognitivist approaches to normativity in anticipation of my later positive account.

In chapters two and three, I look at the nature of naturalism in some detail. Whilst naturalism is very plausibly the dominant metaphilosophical position in contemporary analytic philosophy, it is used in various, sometimes mutually inconsistent, senses. Most notably, we can distinguish ontological and epistemological varieties of naturalism. In chapter two, I argue that the latter

variety is more fundamental. Naturalism is best characterised as a commitment to the view that the natural sciences enjoy a uniquely privileged epistemic status. If the natural sciences are to have this preeminent role, however, we might well wonder what makes them unique. This is closely related to the problem of demarcating science from non-science. Of course, a commitment to epistemological naturalism places constraints on an answer to the demarcation problem. In chapter three, I consider a number of possibilities. Finally, I propose that we take refereed scientific articles as proxy indicators of scientific practice. A study of norms governing scientific publication indicates that an impersonal, objectifying stance is typical of scientific practice.

Having explained the problem of normativity in terms of value-makers, and characterised the naturalistic approach that leads to it, I shall be in a position to examine putative solutions. In chapter four, I consider Penelope Maddy's project of providing scientific foundations for logic. In particular, I look at whether the kinds of facts that make up scientific explanations can be identical to value-makers. I argue, however, that even if Maddy provides an explanation of our practices of logical inference, her account does not thereby contain the resources to justify those practices. I consider whether logical psychologism fares better, and argue that although it is better than its reputation, it also fails.

If scientific facts are not intrinsically normative, they may still have normative significance in virtue of their relational properties. Thus, in chapter five, I consider the popular strategy of reducing normative reasons to pragmatic or instrumental reasons. Here I focus on the well-developed literature on naturalised epistemology and engage with the work of Quine, Stich, Kornblith and Papineau. I argue, however, that instrumentalism with respect to epistemic normativity faces critical difficulties and I highlight a number of implausible results stemming from a conflation of epistemic and pragmatic reasons.

In the second half of the dissertation, I construct a positive account of the value-makers for normative propositions. Chapter six sets the agenda by looking at the relationships between certain key normative concepts. In order to gain a handle on these concepts, I sketch a system of ethics for the relatively simple practice of

chess. This involves distinguishing between the deontological and the evaluative. I associate the former with the necessary preconditions for engaging in valuable activity and the latter with the satisfaction of the ends internal to practices. The account I defend primarily draws on insights from the neo-Aristotelian traditions of virtue ethics and epistemology, but also has certain Kantian themes.

The last three chapters of the dissertation are concerned with developing this account of normativity and applying it beyond the relatively clearly defined and limited practice of chess. In chapter seven, I return to the topic of epistemic normativity. As well as providing an interesting case study, this allows me to pick up and resolve themes from my earlier discussion of instrumentalism. On the evaluative side, I argue that truth, and truth alone, is the aim of belief and consequently the source of epistemic value. I explain how this makes sense of our practices of epistemic evaluation, focussing on virtuous dispositions of thought. As a potential challenge to my account, I consider a thought experiment suggested by Jason Stanley, which intuitively suggests that ascriptions of knowledge depend upon practical interests. This gives me the opportunity to disentangle competing sources of normativity.

In chapters eight and nine, I return to the idea that normative reasons not only justify our actions and beliefs, but also motivate them; when we are rational at least. In chapter eight, I focus on the claim that desires must figure in any naturalistically acceptable account of normative force. I consider the Humean claim that desires are necessary for motivation, but I argue that there is no compelling reason to accept this view. I also argue that a desire-based theory of reasons leaves us with an unattractive anti-rationalism with respect to ends that is incompatible with our practices of evaluation and our moral development.

Finally, chapter nine moves from epistemic normativity to the epistemology of normative properties. Mackie famously argued against moral realism by suggesting that moral properties would have to be ontologically and epistemologically queer. Having already provided an account of what normative properties are, I frame chapter nine in relation to the epistemological challenge.

Firstly, I show how Mackie's epistemological argument depends on assumptions concerning the nature of moral motivation. I then argue that these concerns can be allayed once we understand the role of the emotions in moral epistemology. Drawing on psychological and neurological findings, as well as insights from the phenomenological tradition, I defend the position that cognitive content and felt motivation can come together when emotional experience has normative content as its intentional object.

## Chapter One

### Hume, Moore and Kripke

In this chapter, I look at three influential discussions of the troubled relationship between naturalism and normativity. I begin with Hume's discussion of 'is' and 'ought', before moving on to G. E. Moore's open question argument (OQA) and its close relationship to Kripke's interpretation of Wittgenstein's rule-following considerations. In part, this is intended to serve as an historical introduction to the subject matter of the thesis. The deeper agenda, however, is to uncover the nature of the difficulty philosophical naturalists face in attempting to account for the normative. Consequently, I shall be more concerned with the problems raised than with their solutions at this stage. I shall argue that Hume, Moore and Kripke/Wittgenstein each provide us with an alternative approach to essentially the same underlying problem. Roughly, for now, we can formulate this as follows: 'how, constrained by naturalistic commitments, can one identify those aspects of the world which serve as the truth-makers/truth conditions of normative propositions/assertions?'

#### *1.1 Hume on 'is' and 'ought'*

Hume's famous discussion of 'is' and 'ought' in Book III of the *Treatise* runs as follows:

In every system of morality, which I have hitherto met with, I have always remark'd, that the author proceeds for some time in the ordinary way of reasoning, and establishes the being of a God, or makes observations concerning human affairs; when of a sudden I am surpriz'd to find, that instead of the usual copulations of propositions, is, and is not, I meet with no proposition that is not connected with an ought, or an ought not. This change is imperceptible; but it is, however, of the last consequence. For as this ought, or ought not, expresses some new relation of affirmation, 'tis necessary that it shou'd be observ'd and explain'd; and at the same time that a reason should be given, for what seems altogether inconceivable, how this new relation can be a deduction from others, which are entirely different from it. But as authors do not commonly use this precaution, I shall presume to recommend it to readers; and am persuaded, that this small attention wou'd subvert all the vulgar systems of morality, and let us see, that the distinction of vice and virtue is not founded merely on the relations of objects, nor perceived by reason. (*T* III, 1, I, 469–70)

If there is a received view of this passage, it sees Hume as defending the logical principle that one cannot deductively derive an ‘ought’ from an ‘is’. Following R. M. Hare (1952: 29), this principle is sometimes known as “Hume’s Law”. For example, Jesse Prinz (2007: 1) describes Hume as showing that ‘there is no way to deduce a statement that has prescriptive force (a statement that expresses an unconditional obligation) from statements that are purely descriptive’. Again, William Casebeer describes Hume’s Law as stating that ‘it is impossible to deductively derive an “ought” statement from a set of premises that contain only “is” statements’ (2003: 128). Any attempt to do so is to commit a logical fallacy, sometimes lumped together with Moore’s different putative error of reasoning under the label of the ‘naturalistic fallacy’.<sup>1</sup>

Hume, then, is thought to have shown that there is an inferential gulf separating propositions concerning the way matters actually stand in the world and propositions concerning how they ought to stand. Moreover, this is thought to be a matter of unimpeachable logic. In this vein Charles Pigden writes:

Hume ... is appealing to the conservative character of deduction. He is making a strictly logical point. A valid inference preserves, but does not extend, the truth. The conclusions are contained within the premises. So, if ‘ought’ does not appear among the premises of an alleged inference, but does figure in the conclusion (if, that is, it expresses a new relation or affirmation), then the inference must be invalid. (1989: 130)

Now, it is true that in the syllogistic logic with which Hume was familiar one cannot deduce an ‘ought’ from an ‘is’. Syllogistic logic is conservative in the rough sense that the conclusion is already contained in the premises. One only gets out of a syllogism what one has already put in. So one cannot derive an ‘ought’ from an ‘is’, any more than one can derive the term ‘almond’ from the term ‘avocado’. If an ‘ought’ is derived from a set of premises containing only ‘is’ statements, either a normative term has been smuggled in, for instance a thick moral concept, or the argument is an enthymeme with a suppressed normative premise.<sup>2</sup>

---

<sup>1</sup> For instance, Åsa Wikforss (2001: 219) refers to ‘the famous “naturalistic fallacy” of trying to derive an “ought” from an “is”’.

<sup>2</sup> For instance, one might derive an obligation from a promise (see Searle 1964 for details of how this can be achieved). The derivation depends however on the fact that promises already involve

Matters are more complicated if we turn to modern logic, which is not conservative in the same sense. Here we should say that one cannot non-vacuously derive an ‘ought’ from an ‘is’. The following is a vacuous example employing the principle that everything follows from a contradiction:

P1: It is raining and it is not raining.

Therefore: You ought to carry an umbrella.

Here is another vacuous example by way of the rule of Disjunction-Introduction:

P1: London is the capital of England

Therefore: Either London is the capital of England or one ought to kill the infidel.<sup>3</sup>

There is, however, a clear intuitive sense that there is something wrong or uninteresting about these examples. One reason for this is that the inferences would remain valid even if the normative conclusions were replaced with non-normative conclusions. The pattern of inference does not require a normative conclusion and so reveals nothing about normative propositions in particular. It is clear, in any case, that the prospects of naturalising normativity are unlikely to turn on such considerations. Moreover, it is clear that Hume did not have such examples in mind. If Hume had been making a purely formal point, then it must have been within a syllogistic framework. But, then, the claim that one cannot deduce an ‘ought’ from an ‘is’ is a trivial thought, on a par with the claim that one cannot deduce the term ‘almond’ from the term ‘avocado’. It would then be implausible that ‘this small attention wou’d subvert all the vulgar systems of morality’ or be of ‘the last consequence’. Moreover, Hume himself describes the closely related idea that ‘reason alone can never give rise to any original idea’ as a ‘very obvious principle’ (*TI*, 3, XIV, 157). This much suggests that we should

---

an entanglement of the descriptive and the evaluative. A promise is a speech-act, which, if it is to be distinguished from a mere collection of sounds and gestures, is meaningful in virtue of the normative standards governing the background institution of promising.

<sup>3</sup> Both examples are adapted from Prior (1960). See also Pigden (1989).

look elsewhere if we are to understand the significance of the passage in question.

One reason commentators may have been misled into thinking that Hume is making a point concerning deductive validity is that Hume refers to the question of ‘how this new relation can be a *deduction* from others’ (*T* III, 1, I, 469, emphasis added). We should notice, however, that ‘deduction’ in Hume’s sense is by no means the same as ‘deduction’ in our sense. As Alasdair MacIntyre observes, the word ‘deduction’ appears rarely in Hume’s writings (1959: 460). Where Hume is clearly referring to what we would call ‘deduction’, his preferred term is ‘demonstration’.<sup>4</sup> This is in keeping with standard eighteenth century usage, where ‘deduction’ and its cognates were used to mean ‘inference’ and its cognates, not ‘entailment’. As MacIntyre writes, Hume ‘is not making a point about entailment – for he does not mention it’ (1959: 465).

A deeper source of error, however, is the tendency to read the passage on ‘is’ and ‘ought’ in isolation from Hume’s broader epistemology and philosophy of mind. For example, Pigden (1989: 129) describes the discussion as ‘Hume’s afterthought’, coming as it does at the end of a section devoted to the defence of what we would now call non-cognitivism. Thus, he sees the passage as cut off from the body of Hume’s work. Attempting to understand the discussion out of context is, however, a serious mistake. In order to see this we need to look at Hume’s theory of reference, which will enable us, in turn, to understand what he in fact meant by ‘deduction’.

According to Hume’s theory of reference, then, mental content resembles that which it represents and reasoning involves re-arranging mental content so that it can be seen to instantiate various relations. Following Locke, in substance if not in nomenclature, Hume held that mental content is present in the form of either

---

<sup>4</sup> This does not mean that Hume’s use of ‘demonstration’ is synonymous with our use of ‘deduction’. In particular, Hume uses ‘demonstration’ in such a way that the conclusion of a demonstrative argument is necessarily true. For instance, Hume writes in the *Abstract* that ‘wherever a demonstration takes place, the contrary is impossible and implies a contradiction’ (*A* 650). This might suggest that Hume limited his use of ‘demonstration’ to deductive arguments with necessarily true premises, such as mathematical proofs. For detailed discussion see Owen (1999: 87–91).



impressions or ideas. Impressions, Hume tells us, are ‘all our sensations, passions and emotions as they make their first appearance in the soul’ (*T I*, 1, I, 1). Ideas, on the other hand, are the materials upon which reason, or the ‘imagination’, operates. Today we might speak of sense data and concepts or thoughts respectively. Central to the relationship between the two categories of perception is the so-called ‘Copy Principle’, according to which all of our ideas are copies of corresponding impressions (*T I*, 3, VII, 96).<sup>5</sup> The relationship is causal. An idea is caused by the corresponding impression, from which it derives its content. Impressions, in turn, derive their content from whatever caused them.<sup>6</sup>

Ideas are importantly different from the propositional content favoured by more recent theories of mind. Although ideas can be propositional in the sense that they can be judgements or beliefs, they do not have the sort of propositional structure which would allow them to enter into inferential relations in the same manner as propositions.<sup>7</sup> Instead of depending on the logical constants which make up the connective tissue of propositional logic, Humean reasoning involves the mind passing from one idea to another via ‘philosophical relations’ (*T I*, 1, V, 14).<sup>8</sup> There are seven philosophical relations, supposedly exhaustive. These are resemblance, identity, space and time, quantity or number, quality or degree, contrariety, and cause and effect. Philosophical relations are an ‘arbitrary union

---

<sup>5</sup> More precisely, all of our simple ideas, ideas that are indivisible in thought, are copies of whichever simple impressions originally caused them. Complex ideas with no correspondent impression, such as a unicorn or a golden mountain, can be constructed from simple ideas.

<sup>6</sup> For Hume, the cause of an impression is necessarily obscure: ‘As to those impressions, which arise from the senses, their ultimate cause is, in my opinion, perfectly inexplicable by human reason, and ‘twill always be impossible to decide with certainty, whether they arise immediately from the object, or are produc’d by the creative power of the mind, or are deriv’d from the author of our being’ (*T I*, 3, V, 84). Metaphysical investigation of ultimate causes violates Hume’s methodological strictures.

<sup>7</sup> Beliefs are distinguished from ideas in terms of their greater ‘force and vivacity’ (*T I*, 3, VII, 96). In this sense, beliefs are intermediate between impressions and ideas.

<sup>8</sup> The mind can also move from one idea to another via principles of association or ‘natural relations’. According to Hume’s psychological theory of associationism, by which he set great store, there are precisely three ways in which the mind ‘is convey’d from one idea to another’ (*T I*, 1, IV, 11), namely resemblance, contiguity in space or time, and cause and effect. Apparently inviting a comparison with Newton’s theory of gravitation, Hume describes association as ‘a kind of ATTRACTION, which in the mental world will be found to have as extraordinary effects as in the natural, and to shew itself in as many and as various forms’ (*T I*, 1, IV, 12–3). Patterns of thought produced by association are structurally analogous to patterns of reasoning, but distinct insofar as they are psychologically determined (see *T I*, 3, VI, 92). For instance, Hume invokes natural relations in order to account for the phenomenology of causal expectation, which cannot, he argues, be grounded in reason.

of two ideas' (*TI*, 1, V, 13) in the sense that any two ideas with which the mind is already stocked can be compared in the imagination via the intermediary of a philosophical relation. However, there is no natural compulsion to reason from one to the other. We reason from one idea to another when we judge that they stand to one another in a particular philosophical relation. For Hume, unlike Descartes, this is always a fallible process.

Although we are free to reconstruct his arguments in terms of premises, conclusions and formal validity, we can already see that this is not how Hume would have envisaged the inference from an 'is' to an 'ought'. Importantly, valid demonstrative reasoning depends on content, rather than form.<sup>9</sup> As we have already mentioned, the content of an idea is causally derived from the corresponding impression. Following from the Copy Principle, an idea resembles the impression that caused it. Hence, if we consider visual impressions as the paradigm case, ideas are imagistic.<sup>10</sup> Following Millgram, we might call Hume's a 'pictorial theory of content' (2005: 220). Demonstrative reasoning, then, requires memory to call up ideas and imagination to rearrange these images in our mind, place them in juxtaposition, and notice the philosophical relations which they bear to one another, either directly or through a chain of intermediate ideas.<sup>11</sup> Thus, Harrison describes mathematical reasoning on the Humean model:

Two and two are four – an a priori necessary truth, discoverable by reason – can be known to be true by comparing our idea of, say, two spots and another idea of two spots, and seeing that they must be equal in number to our idea of four spots. (1976: viii)

---

<sup>9</sup> See further Owen (1999: 3).

<sup>10</sup> There are, of course, ideas corresponding to impressions from non-visual modalities. Owen suggests with some misgiving that we might think of Hume's example of the idea of the taste of a pineapple as a 'gustatory image of an impression of taste' (1999: 73). Although this usage tends to jar, committing us to non-visual images, it does at least help us keep in mind the fact of resemblance between impressions and ideas. It seems easier to understand how an imagistic idea can resemble a visual impression than, say, how an idea of a taste resembles the corresponding impression. Perhaps this is due to our greater facility at isolating, manipulating and simultaneously comparing visual images than smells, sounds, feels or tastes.

<sup>11</sup> 'All kinds of reasoning consist in nothing but a comparison, and a discovery of those relations, either constant or inconstant, which two objects bear to each other' (*TI*, 3, II, 73). Although the 'objects' Hume mentions can be either ideas or impressions, if both objects are impressions it is a case of perception rather than reasoning.

In this case, the relevant philosophical relation is quantity or number.<sup>12</sup>

So, on this model of reasoning, inferring an 'ought' from an 'is' would involve bringing to mind both the idea of a particular action and the idea of its appropriateness, and noticing how the two ideas are related. The manner in which the two ideas are related will depend on their respective content, which will in turn depend on their origins.<sup>13</sup> As the content of an idea derives from the content of the corresponding impression, Hume's theory of reference also provides a method for identifying bogus or meaningless ideas:

When we entertain, therefore, any suspicion that a philosophical term is employed without any meaning or idea (as is but too frequent), we need to enquire, from what impression is that supposed idea derived? And if it be impossible to assign any, this will serve to confirm our suspicion. (*E* 22; see also *E* 74, *A* 648–9)

Whenever there is no corresponding impression, an idea is devoid of content. Or, more strictly, it is devoid of the content we previously took it to have. For example, we naively take ourselves to have an idea of necessary connection. However, careful empirical investigation reveals that there is no impression from which this idea could derive its content. Rather, we have numerous impressions of the regular conjunction of two events. Hence, it turns out that our idea of necessary connection is actually an idea of constant conjunction. Strictly, therefore, we have no idea of necessary connection, merely the appearance of one.<sup>14</sup> This is perhaps a surprising result, entailing that empirical investigation can correct errors concerning the contents of our own minds. More pertinently

---

<sup>12</sup> Millgram (2005: 220) observes that 'in Hume's day the foremost deductive science was geometry, in which the reasoning was explicitly pictorial'. An emphasis on geometric as opposed to algebraic proofs increases the intuitive appeal of an imagistic account of mathematical reasoning.

<sup>13</sup> Hume writes as follows: 'Tis impossible to reason justly, without understanding perfectly the idea concerning which we reason; and 'tis impossible perfectly to understand any idea, without tracing it up to its origin, and examining that primary impression, from which it arises. (*TI*, 3, II, 74–5)

<sup>14</sup> Drawing a distinction between the appearance and reality of an idea is obviously not without difficulty. It may be for this reason that Baillie (2000: 28) claims that Hume '[does not] deny that we have an idea of this necessary connection'. Here, however, is what Hume actually writes: 'Every idea is copied from some preceding impression or sentiment; and where we cannot find any impression, we may be certain that there is no idea' (*E* 78). Again, 'when we speak of a necessary connection betwixt objects, and suppose that this connection depends upon an efficacy or energy, with which any of these objects are endow'd; in all these expressions, so apply'd, we have really no distinct meaning, and make use only of common words, without any clear and determinate ideas' (*TI*, 3, XIV, 162).

for present purposes, it shows that an idea ungrounded in an appropriate impression is without its apparent content. As Humean reasoning relies on content rather than form, such an idea cannot moreover enter into inferential relations (or not those which we originally thought). Hence, questions concerning the possibility of reasoning involving normative ideas essentially depend upon how and if the ideas are associated with impressions. What, then, is the source of our normative ideas?

It is here that Hume's empiricism is significant. It was Hume's view that a tradition of speculative metaphysics, reasoning from uncertain first principles, had 'drawn disgrace upon philosophy itself' (T intro, xiii). Impressed above all by Newton and the explanatory successes of the natural sciences, Hume sought to place philosophy on an equally sound scientific footing by introducing similarly strict methodological discipline. This project is clearly indicated by the subtitle of the *Treatise*, 'being an attempt to introduce the experimental method of reasoning into moral subjects'. Hume writes:

And tho' we must endeavour to render all our principles as universal as possible, by tracing up our experiments to the utmost, and explaining all effects from the fewest and simplest causes, 'tis still certain we cannot go beyond experience; and any hypothesis, that pretends to discover the ultimate original qualities of human nature, ought at first to be rejected as presumptuous and chimerical. (*T*, intro, xvii)

The impressions that cause our normative ideas must therefore be located via experience. The difficulty is that there is no clear candidate for an empirically discoverable property which resembles our idea of the appropriateness of an action distinct from the action itself:

Take any action allow'd to be vicious: Wilful murder, for instance. Examine it in all lights, and see if you can find that matter of fact, or real existence, which you call vice. In which-ever way you take it, you find only certain passions, motives, volitions and thoughts. There is no other matter of fact in the case. The vice entirely escapes you, as long as you consider the object. (*T* III, 1, I; 468)<sup>15</sup>

---

<sup>15</sup> There are clear parallels here with Hume's discussion of 'necessary connexion' in the context of causation: 'Here again I turn the object on all sides, in order to discover the nature of this necessary connexion, and find the impression, or impressions, from which its idea may be deriv'd. When I cast my eye on the known qualities of objects, I immediately discover that the

Ideas derive their content from imagistic impressions, but there is no imagistic difference between the representation of an action and the representation of the appropriateness or inappropriateness of an action. Millgram expresses this point clearly:

The easiest way to see this is to imagine trying to use a picture of a state of affairs to represent the necessity or obligatoriness of that state of affairs by modifying the representational elements of the picture – perhaps by scrawling “Necessary” or “Obligatory” across the top. The attempt is bound to fail: what one will get is not a picture of necessity, but a picture of, say, a landscape marred by peculiar skywriting. (2005: 235)

If, however, there is no difference between the two images, then there is no difference in content which would mark a distinction between normative and non-normative ideas. The conclusion is that if our normative ideas have specifically normative content this must derive from some source other than our experience of the external world.<sup>16</sup>

At this stage, one might be inclined to maintain that we do perceive normative properties in the world in addition to natural properties. In order to pursue this line of argument one might object to Hume’s pared down phenomenology. Isn’t it the case, one might ask, that we see an action as virtuous or vicious, that we are directly aware of the goodness or wickedness of an action and that this awareness is intrinsic to our phenomenological experience? It is far from easy, however, to flesh out this intuitively attractive claim in a way that does not violate Hume’s methodological principles. In particular, Hume is constrained by his account of the nature of perception, according to which ‘my senses convey to

---

relation of cause and effect depends not in the least on them. When I consider their relations, I can find none but those of contiguity and succession’ (*TI*, 3, II, 77).

<sup>16</sup> Not only shall we fail to find normative properties by studying actions or states of affairs in themselves, but also, Hume argues, we shall fail if we consider the relations holding between them. He employs the following argument: ‘Let us chuse any inanimate object, such as an oak or elm; and let us suppose, that by the dropping of its seed, it produces a sapling below it, which springing up by degrees, at last overtops and destroys the parent tree: I ask, if in this instance there be wanting any relation, which is discoverable in parricide or ingratitude?’ (*T III*, 1, I, 467). Our normative ideas cannot be grounded in a relation because two states of affairs can instantiate the same relations, but differ in terms of their normative properties. One might argue that the relation holding between an oak and an acorn is quite different from that between human relatives insofar as the nature of the relation depends on the nature of the relata. The challenge inherent in this line of response is either to specify the nature of the difference in non-normative terms, so as to avoid charges of circularity, or to show that the circularity is not vicious. For a defence of the latter approach see Millgram (2005: 227–8).

me only the impression of colour'd points, dispos'd in a certain manner' (*T* I, 2, III, 34). Light does not seem to carry normative information and it is hard to see how other modalities, accounted for in mechanistic terms, would be better equipped to do so. There is no obvious candidate for, say, the smell of goodness or the taste of evil. An appeal to a suprasensible faculty of intuition obviously would not do for Hume.<sup>17</sup>

Thus, Hume concludes that the external world, as revealed by empirical experience, affords us no impressions which could ground our normative ideas. Hence, the inference from 'is' to 'ought' cannot proceed if normative ideas are supposed to derive their content from corresponding external normative impressions.<sup>18</sup> There are no such impressions and, so, normative ideas would be meaningless and unsuitable for entering into inferential relations. Although the point, thus expressed, concerns meaning, we have seen that the root of the problem is epistemological, rather than logico-semantic in the modern sense. How, from a particular epistemic standpoint, subject to certain methodological restrictions, are we to locate the referent of normative properties or relations? As Hume has a causal theory of reference and the investigation of causal relations apparently lies within the proper domain of the natural sciences, we might gloss the question as 'how can science identify normative properties?'

---

<sup>17</sup> My central claim is of course that Hume's problem of locating the referents of normative ideas stems from his naturalism. However, it is interesting to note an alternative line of argument, according to which Hume's problem might result from an insufficiently naturalistic approach. We can extract the argument from certain thoughts of John Dewey, starting from the idea that in scientific method 'the vine of pendant theory is attached at both ends to the pillars of observed subject-matter' (1929: 2a). Hume's model of the relationship between mind and world is of Newtonian inspiration. Newtonian physics, grounded first in experience, developed in accordance with reason, returns to the second pillar of observed subject matter in experiment and confirmation. Hume adopts Newton's mechanistic model of the world and, in this light, develops his account of concept acquisition. However, the vine is left swinging. Rather than testing his model against experience, Hume interprets his experience in light of his theoretical commitments, with the arguably implausible result that his experience is claimed to be devoid of normative qualities. Dewey might have been referring to Hume when he wrote, 'The discoveries and methods of the physical science, the concepts of mass, space, motion, have been adopted wholesale in isolation by philosophers in such a way as to make dubious and even incredible the reality of the affections, purposes and enjoyments of concrete experience.' (1929: 35) Dewey, himself an empirical naturalist, does not deny the reality of the value in the world as revealed in experience.

<sup>18</sup> Notice how the two tines of Hume's infamous fork are related in this line of argument. It is because putative normative facts are not matters of fact that they cannot be inferred from non-normative facts as relations of ideas.

Hume, of course, did not finally doubt the meaningfulness of moral discourse. Famously, he maintained that the source of our ideas of morality is available upon introspection:

You never can find it [vice], till you turn your reflexion into your own breast, and find a sentiment of disapprobation, which arises in you, towards this action. Here is a matter of fact; but 'tis the object of feeling, not of reason. It lies in yourself, not in the object. So that when you pronounce any action or character to be vicious, you mean nothing, but that from the constitution of your nature you have a feeling or sentiment of blame from the contemplation of it. (*T* III, 1, I; 468–9)

A normative idea derives its content from the passions, from a 'feeling or sentiment'. In order to distinguish our normative ideas from our non-normative ideas, a further element is adjoined to the image, not a further image, but a feeling of approbation or disapprobation. Hume calls the passions 'secondary' or 'reflective' impressions, in order to illustrate that they arise in response to 'original' impressions, either of the external world or of internal sensations (*T* II, 1, I, 275). Our passions resemble our ideas of virtue and vice insofar as they have a felt aspect of approbation or disapprobation and corresponding motivational force.<sup>19</sup>

Although the object of approbation or disapprobation is the ultimate cause of the normative idea, it is not the source of its content. The situation is analogous to the case in which our perception of one event constantly conjoined to another gives rise to the idea of necessary connection, the content of which, however, can neither be derived from the events considered in themselves nor from the relation between the two events. Rather, we must 'enlarge [our] view to comprehend several instances; where [we] find like objects always existing in like relations of contiguity and succession' (*T* I, 3, XIV, 155). This repetition gives rise to a new impression of 'determination', from which the content of his idea of necessary connection derives. An important consequence is that our idea of necessary connection turns out not to be about its putative objects. Similarly, our ideas of

---

<sup>19</sup> The association between the passions and the affective elements of our moral psychology remains at the heart of contemporary arguments in favour of non-cognitivism. Hume's claim that 'reason alone can never be a motive to any action of the will' (*T* II, 3, III, 413) remains influential. We shall return to this later in the chapter.

virtue and vice turn out not to be about their putative objects, although we have a natural tendency to project our normative ideas onto the external world.

The content of our normative ideas, deriving from internal passions, is not, then, wholly pictorial and does not resemble its apparent object. As the content of normative ideas is unlike the content of ideas derived from sensory impressions, we cannot reason from one to the other via philosophical relations.<sup>20</sup> Although the normative idea arises in response to its apparent object (either directly in response to the impression, or in response to the subsequent idea), they are logically and psychologically independent.<sup>21</sup> There is no necessary connection between any given impression and a feeling of approbation or disapprobation. This is the line of thought underpinning the following striking passage:

‘Tis not contrary to reason to prefer the destruction of the whole world to the scratching of my little finger. ‘Tis not contrary to reason for me to chuse my total ruin, to prevent the least uneasiness of an Indian or person wholly unknown to me. ‘Tis as little contrary to reason to prefer even my own acknowledg’d lesser good to my greater, and to have a more ardent affection for the former than the latter. (*T* II, 3, III, 415–6)

I have now dealt at some length with the details of Hume’s theory of reference and his account of demonstrative reasoning. This seemed necessary for several reasons. First, there is the intrinsic value of a correct understanding of Hume’s discussion of ‘is’ and ‘ought’. Second is the influence of Hume’s argument. Millgram claims that ‘today [Hume’s Law] is a philosophical near-truism, and the burden of proof is taken to rest squarely on the shoulders of its opponents’ (2005: 218). If he exaggerates, it is not by much. Hume’s discussion has, moreover, filtered out from academic philosophy and established itself as received wisdom across swathes of the natural and social sciences. For example, we read the following in a recent *Nature* editorial:

---

<sup>20</sup> Of the seven philosophical relations Hume considers only four, ‘resemblance, contrariety, degrees in quality, and proportions in quality and number’ (*T* III, 1, 1, 464, emphasis removed) as putative candidates. Although Hume does not explicitly explain this constraint, it is straightforward to see what motivates it. Hume holds identity to be more properly a matter of perception than of reasoning (*T* I, 3, II, 74), and cause and effect, and space and time are more properly associated with probabilistic reasoning concerning matters of fact; they hold between external objects rather than between an internal action (that of approbation or disapprobation – an ‘ought’) and an external object (an ‘is’).

<sup>21</sup> As Hume subscribes to logical psychologism (*T*, intro, xv), this finally reduces to the same point.



It becomes particularly hard – and at the same time especially important – to resist the ‘naturalistic fallacy’ of inferring what ought to be from what is. Science may be able to tell us why some values are more easily held than others. But it cannot tell us whether taking the easy path in terms of which values we espouse is the right thing to do’ (2009: 763).

No defence of the status of the ‘naturalistic fallacy’ as a fallacy is offered. Nor, indeed, would any be expected by an educated readership. If, however, such claims owe an implicit debt to Hume’s authority, it is of practical importance that we understand Hume’s argument correctly. Third and most important for present purposes, a faulty understanding of Hume’s discussion blinds us to its proper significance. Were Hume making a point concerning formal validity in the modern sense, we might be inclined to join Richard Joyce in saying ‘Who cares?’ (2006: 153). This would be unfortunate as Hume’s argument, when it is properly understood, remains vital to concerns about the relationship between the natural and the normative. It does so by providing a clear example of the manner in which an empirical stance can render the idea of realistically construed mind-independent norms mysterious or, even, meaningless. I shall return to this problem at the conclusion of the chapter. First, I propose to turn to two further arguments, both of which can be plausibly interpreted as purporting to show the inadequacy of the natural facts to determine the truth of normative proposition. These are Moore’s OQA and the ‘sceptical paradox’ which Kripke finds in the writings of Wittgenstein. I shall consider each in turn.

### *1.2 Moore’s Open Question Argument*

Moore argued that any proposed definition of ‘good’ commits the ‘naturalistic fallacy’.<sup>22</sup> Unfortunately, the label is peculiarly opaque insofar as the alleged error is neither a logical fallacy nor does it exclusively apply to naturalistic definitions of ‘good’. The naturalistic fallacy is intended to operate equally against ‘metaphysical’ definitions, asserting an identity between ‘good’ and

---

<sup>22</sup> By ‘definition’, Moore means something like a real definition, identifying what goodness is with some other property, rather than a nominal definition. Thus, he writes that ‘verbal questions are best left to the writers of dictionaries and other persons interested in literature; philosophy, as we shall see, has no concern with them’ (PE 2). Note also that Moore does not object to definitions of ‘the good’, as opposed to ‘good’, if by defining ‘the good’ we mean establishing its extension. This amounts to the claim that we are entitled to say ‘X is good’ only when using the ‘is’ of predication, as opposed to the ‘is’ of identity.

some object existing in a ‘supersensible real world’ (*PE* 39). Nevertheless, we shall see that Moore’s OQA does pose a challenge to naturalistic ethics.

Before outlining the structure of the argument in its most familiar form, it is worth noting that, as with Hume’s discussion of ‘is’ and ‘ought’, the received view of Moore’s OQA cannot be supported by a close examination of the text. Although I spent considerable time unravelling the historical Hume’s line of argument, I shall not pursue the same course with respect to Moore. Hume’s discussion turned out to be stronger and more interesting than its standard interpretation. The reverse is true of Moore’s OQA. As Moore himself later came to recognise, his presentation is muddled and fails in its aims (1932: 127). Not only is his central line of argument weak, but he also vacillates concerning the precise nature of the alleged fallacy. Competing passages suggest any of the following interpretations: that to commit the naturalistic fallacy is (i) to define goodness, (ii) to assert of some property other than goodness that it is goodness, (iii) to deny that goodness is a simple property or (iv) to deny that goodness is non-natural. Thomas Baldwin appropriately concludes that ‘Moore’s discussion is hopelessly confused on this matter’ (1990: 70). Although these claims require substantiation, an excursion into detailed exegesis would tend to divert us from our main topic. Moreover, the matter has been discussed perfectly well elsewhere.<sup>23</sup> These brief comments, then, are merely by way of introducing the caveat that whilst I take *Principia Ethica* as my primary text, my aim is a rational reconstruction rather than an interpretation of the OQA. I am more interested in the insight driving the OQA than in its faulty expression.

As it is commonly understood, then, the OQA runs as follows (*PE* 15–6). Take any proposed definition of ‘good’. Moore borrows Russell’s example of ‘what we desire to desire’. Next, consider some property A, which we judge to be good. Now, frame the following questions:

(1) A is good, but is it what we desire to desire?

---

<sup>23</sup> See Baldwin (1990: 69–73).

(2) A is what we desire to desire, but is it good?

Given the postulated synonymy between ‘good’ and ‘what we desire to desire’, the above questions should appear to be of the same sort as the following trivial questions:

(3) A is something we desire to desire, but do we desire to desire it?

(4) A is good, but is it good?

It is evident, however, that (1) and (2) differ from (3) and (4). Whilst we cannot answer (3) or (4) in the negative without conceptual confusion, (1) and (2) appear significant and undecided in advance. Whilst (1) and (2) are open questions, (3) and (4) are closed. Moore supposed this to show that ‘good’ cannot simply mean ‘what we desire to desire’:

“That we should desire to desire A is good” is not merely equivalent to “That A should be good is good.” It may indeed be true that what we desire to desire is always also good; perhaps, even the converse may be true: but it is very doubtful whether this is the case, and the mere fact that we understand very well what is meant by doubting it, shews clearly that we have two different notions before our minds. (*PE* 16)

Although one would have to proceed by enumeration, the argument appears to generalise against any definition.<sup>24</sup> Moore concluded that the property of goodness must be simple, unanalysable, and non-natural.

Whilst this compact argument has been widely discussed over the past century or so, it receives little contemporary support. Briefly, Moore appears to have

---

<sup>24</sup> Note, however, that there is reason to question the validity of this move. Quite obviously, the handful of examples Moore actually considers constitute a slim basis for concluding that *no* natural property is coreferential with ‘good’. This may suggest that the appeal to generalisation is motivated by some independent principle. One possibility is the definist fallacy discussed below. Sturgeon (2003: 536) raises a further objection. He notes that there *is* one property which we can be perfectly certain is coreferential with ‘good’. Trivially, ‘good’ is coreferential with itself. So, Moore’s argument that ‘good’ is not coreferential with any natural property only goes through on the assumption that ‘good’ does not itself refer to a natural property. In other words, the argument begs the question against the Cornell Realist view that goodness is a *sui generis* natural property. This does not, however, show that the OQA is similarly question-begging with respect to reductive naturalistic proposals. For we cannot be sure in advance that the proposed reduction will not be vulnerable to the OQA.

overlooked that asking is an opaque context. In Fregean terms, even if ‘good’ and ‘what we desire to desire’ have different senses, they might, for all the argument shows, have the same reference. For instance, one can ask whether the ‘morning star’ is the ‘evening star’ without betraying a lack of conceptual or linguistic competence. This is because ‘morning star’ and ‘evening star’ have different senses. Empirical investigation, nevertheless, reveals them to have the same reference, i.e. Venus. According to Kripke’s influential account this is, indeed, a necessary truth. Moreover, identity statements apparently susceptible to the OQA, but in fact true, are not limited to the a posteriori. One reason is conceptual difficulty. The proposition ‘ $e^{i\pi} + 1 = 0$ ’ is less likely to command universal assent than ‘ $0 = 0$ ’. Minor computational difficulty suffices; compare ‘ $99^2 = 9801$ ’ with ‘ $9801 = 9801$ ’. And, if mathematical complexity can undermine introspective tests of identity, there is little reason to suppose that ethics is an easier case. Thus, the OQA as it stands neither rules out a broadly naturalistic account of normativity, nor rules in a posteriori methodology as the only viable means of generating such an account.

Before considering how the OQA can be strengthened, we should notice a further reason to be sceptical concerning its significance to the project of naturalising normativity, namely that it is unclear whether the OQA has any special relevance to the normative. Thus, it is often observed that Moore’s OQA is not directed against definitions of moral terms such as ‘good’ in particular, but targets definitions of simple properties quite generally; colour properties such as ‘yellow’, for instance.<sup>25</sup> William Frankena argues that the naturalistic fallacy is but an instance of a more general ‘definist fallacy’:

The definist fallacy is the process of confusing or identifying two properties, of defining one property by another, or of substituting one property for another. Furthermore, the fallacy is always simply that two properties are treated as one, and it is irrelevant, if it be the case, that one of them is natural or non-ethical and the other non-natural or ethical. One may commit the definist fallacy without infringing on the bifurcation of the ethical and the non-ethical, as when one identifies pleasantness and redness or rightness and goodness. (1939: 471)

---

<sup>25</sup> E.g. Joyce (2006: 148) and Bloomfield (2006: 172–3).

As Frankena observes, this fallacy merely expresses Bishop Butler's maxim, quoted with approval by Moore opposite the title page of *Principia Ethica*, that 'everything is what it is, and not another thing'.<sup>26</sup> If this is so, Moore's naturalistic fallacy bears no special affinity to Hume's problem, or to the central theme of the present thesis. Furthermore, the definist fallacy threatens to turn into universal acid in the form of the paradox of analysis. It then becomes difficult to resist the following view advanced by Darwall, Gibbard and Railton:

To grant Moore all of the resources he deploys or assumes in his official presentation of the open question argument would suffice to bring the whole enterprise of conceptual analysis to a standstill and show nothing about Good in particular. (1992: 3)

Frankena's interpretation might be correct, although it is difficult to reconcile with Moore's insistence that there is something special about definitions of goodness, as opposed to other properties, which merits the label 'naturalistic fallacy'.<sup>27</sup> Certainly, passages can be found in Frankena's defence.<sup>28</sup> The case is difficult to adjudicate, for, as mentioned above, Moore equivocates concerning the precise nature of the fallacy. In any case, our present concern is with the normative and it is here that the OQA appears most interesting, and, perhaps, most potent.

Departing now from Moore's text, let us consider the following question: 'What accounts for the intuitive openness of putative identifications of normative properties with natural properties?' It has struck a number of philosophers that

---

<sup>26</sup> See also *PE* 206.

<sup>27</sup> I have in mind the following passage: 'When a man confuses two natural objects with one another, defining the one by the other, if for instance, he confuses himself, who is one natural object, with "pleased" or with "pleasure" which are others, then there is no reason to call the fallacy naturalistic. But if he confuses "good," which is not in the same sense a natural object, with any natural object whatever, then there is a reason for calling that a naturalistic fallacy; its being made with respect to "good" marks it as something quite specific, and this specific mistake deserves a name because it is so common' (*PE* 13). The difficulty is that Moore never explains why it is that good marks out something quite specific. My following discussion is, in part, an attempt to fill this lacuna.

<sup>28</sup> For example, Moore concludes the chapter entitled 'Naturalistic Ethics' as follows: 'In this chapter I have begun the criticism of certain ethical views, which seem to owe their influence mainly to the naturalistic fallacy – the fallacy which consists in identifying the simple notion which we mean by 'good' with some other notion' (*PE* 58). Against this interpretation, however, we should note that Moore himself goes on to define beauty: 'it appears probable that the beautiful should be defined as that of which the admiring contemplation is good in itself' (*PE* 201). It is difficult to see how Moore can be rendered consistent on this point.

the best candidate explanation is the fact that normative ideas motivate us to act.<sup>29</sup> For instance, Darwall, Gibbard and Railton write:

Attributions of goodness appear to have a conceptual link with the guidance of action, a link exploited whenever we gloss the open question “Is P really good?” as “Is it clear that, other things equal, we really ought to, or must, devote ourselves to bringing about P?” (1992: 4)

As we saw above, it is the fact of the intimate connection between normative ideas and motivation that Hume relied upon in order to distinguish our ideas of objects from our normative ideas, which we naively take to be about those objects. Judging that some action or state of affairs is morally right obligates us and, minimally, typically motivates the virtuous among us to act accordingly. Judging that some action or state of affairs instantiates certain natural properties, however, is compatible with indifference. This was Hume’s point in maintaining that “‘Tis not contrary to reason to prefer the destruction of the whole world to the scratching of my little finger’ (*T* II, 3, III, 415).

Although the OQA is directed in the first place against the moral property of goodness, the argument applies *mutatis mutandis* to other normative properties, moral and non-moral, insofar as normative properties are associated with motivation to act (and, more controversially, to believe). Perhaps the most popular view is that the recognition of goodness causes, or is otherwise associated with, an affective mental state such as a desire, which necessarily motivates. It is this line of thought, which, contrary to Moore’s intentions, led to non-cognitivism as the dominant line of response to the OQA in the twentieth century. However, we have no current need to make this assumption or to take sides in the ongoing debates concerning internalism. Let us instead label the motivational aspect ‘normative force’, leaving aside its nature and origins for the present. We reach the conclusion, then, that it is the normative force associated

---

<sup>29</sup> E.g. Hare (1952: 30), Baldwin (1990: 89–90), Darwall, Gibbard and Railton (1992: 4–5), Jackson (1998: 153) and Bloomfield (2006: 170). Blackburn sums of the idea as follows: ‘The reason expressivism in ethics has to be correct is that if we supposed that belief, denial, and so on were simply discussions of a way the world is, we would still face the open question. Even if that belief were settled, there would still be issues of what importance to give it, what to do, and all the rest. For we have no conception of a “truth condition” or fact of which mere apprehension by itself determines practical issues. For any fact, there is the question of what to do about it. But evaluative discussion just is discussion of what to do about things’ (1998: 70).

with the idea of goodness which keeps open questions open. Complete confidence in identifying goodness with some natural property is precluded by the fact that we can reasonably question whether knowledge of the natural property's instantiation suitably motivates us.

Before moving on to look at the rule-following considerations, it is worth emphasising a further similarity between Moore and Hume. The likeness emerges most clearly when Moore recasts his argument in primarily metaphysical terms:

Immensely the commonest type of truth, then, is one which asserts a relation between two existing things. Ethical truths are immediately felt not to conform to this type, and the naturalistic fallacy arises from the attempt to make out that, in some roundabout way, they do conform to it. It is immediately obvious that when we see a thing to be good, its goodness is not a property which we can take up in our hands, or separate from it even by the most delicate scientific instruments, and transfer to something else. (*PE* 124)

Later, in summing up this discussion, Moore writes that 'here, then, we have the root of the naturalistic fallacy' (*PE* 125). The project of naturalising normativity involves asserting an identity relation between some natural object and some normative property. Natural objects are those objects which, in the first place, we can 'take up in our hands', or the properties of such objects as can be detected by scientific instruments. Although tangibility is a rough and ready criterion, the thought seems to follow from Moore's view of the natural as that which exists in time (*PE* 40).<sup>30</sup> The property of goodness, by contrast, does not appear to be the sort of property which can be handled or probed with scientific instruments. Scientific methodology is, therefore, inadequate to the task of providing knowledge of normative properties such as goodness. Thus, Moore argues, we

---

<sup>30</sup> Moore is commonly represented as endorsing a methodological account of the natural (e.g. Miller 2003), writing that 'by "nature" then, I do mean and have meant that which is the subject-matter of the natural sciences and also of psychology' (*PE* 40). However, he also writes: 'If we consider whether any object is of such a nature that it may be said to exist now, to have existed, or to be about to exist, then we may know that that object is a natural object, and that nothing, of which this is not true, is a natural object' (*PE* 40). This suggests that Moore's understanding of the natural is more properly metaphysical than methodological. The natural is that which exists in time, this metaphysical property determining the scope of the natural sciences (and psychology). Hume's naturalism, as we saw above, amounts to methodological strictures. Constrained in this way, he holds aloof from metaphysical speculation concerning the causes of our impressions. I shall spend considerable time analysing these two types of naturalism in the next chapter.

must rely on the cognitive faculty of intuition. The argument reconstructed in this way returns us to the Humean scientific investigator surveying the world, discerning facts but no values, and concluding that the two cannot be identical.

### 1.3 The Rule-Following Considerations

The ‘sceptical paradox’ that Kripke (1982) found in the writings of Wittgenstein arises from the following line of argument.<sup>31</sup> Suppose that we come to calculate the sum ‘68 + 57’ for the first time. Let us suppose that we have previously added only smaller numbers.<sup>32</sup> Given a correct understanding of ‘+’, we should of course answer that ‘68 + 57 = 125’. Kripke, however, looks to shake our confidence by confronting us with a Wittgensteinian challenger:

Perhaps, he suggests, as I used the term ‘plus’ in the past, the answer I intended for ‘68 + 57’ should have been ‘5’! ... After all, he says, if I am now so confident that, as I used the symbol ‘+’, my intention was that ‘68 + 57’ should turn out to denote 125, this cannot be because I explicitly gave myself instructions that 125 is the result of performing the addition in this particular instance. By hypothesis, I did no such thing. But, of course, the idea is that, in this new instance, I should apply the very same function or rule that I applied so many times in the past. But who is to say what function this was? In the past I gave myself only a finite number of examples instantiating this function. All, we have supposed, involved numbers smaller than 57. So perhaps in the past I used ‘plus’ and ‘+’ to denote a function which I will call ‘quus’ and symbolize by ‘⊕’. It is defined by

$$\begin{aligned} x \oplus y &= x + y, \text{ if } x, y < 57 \\ &= 5 \text{ otherwise} \end{aligned}$$

---

<sup>31</sup> There is considerable scope for debate as to whether, or to what extent, Kripke’s Wittgenstein (KW) corresponds to the historical Wittgenstein. Colin McGinn, for example, maintains that ‘what Kripke has done is to produce an impressive and challenging argument which bears little affinity with Wittgenstein’s own problems and claims: in an important sense Kripke and the real Wittgenstein are not even dealing with the same issues’ (1984: 60). Other philosophers who doubt that Kripke has the correct interpretation of Wittgenstein include Blackburn ([1984] 2002) and McDowell ([1984] 2002). One reason for concern is that an acceptance of the ‘sceptical paradox’ is difficult to reconcile with Wittgenstein’s professed ‘quietism’. Martin Kusch, by contrast, argues that ‘Kripke’s interpretation of the sections on rule-following in *PI* and the *Remarks on the Foundations of Mathematics* is, by and large, on target’ (2006: 237). The matter is complicated further as Kripke disassociates himself from the arguments he presents, ‘with few exceptions’ (1982: 5). For present purposes, I am officially content to attribute the arguments to the fictional character KW. In practice, this means that I shall take Kripke (1982) as my primary text, referring to Wittgenstein only in order to clarify or magnify certain aspects of the argument.

<sup>32</sup> In fact, insisting that one has not previously performed the addition is unnecessary, although helpful to get the paradox off the ground. For as Wittgenstein observes in *Remarks on the Foundations of Mathematics*, ‘If I know it in advance, what use is this knowledge to me later on? I mean: how do I know what to do with this earlier knowledge when the step is actually taken?’ (1978: 1–3; see also Kripke 1982: 52n).



Who is to say that this is not the function I previously meant by ‘+’?  
(1982: 8-9)

The sceptical challenge is to cite some fact or facts that suffice to show that we previously meant addition rather than quaddition. Moreover, the facts we cite must enable us to read off what constitutes correct and incorrect usage of ‘+’. It is not enough, that is, to show that we actually use ‘+’ in a particular way. We must also show why we ought to use ‘+’ in the way we do. Only then will the sceptic be satisfied that we are justified in responding ‘125’, not merely employing signs arbitrarily. The sceptic argues that this challenge cannot be met, and so there is no fact of the matter in virtue of which sentences such as ‘Jones meant addition by “+”’ are true. Furthermore, the arguments leading to scepticism concerning the meaning of ‘+’ can be directed towards meanings quite generally, with the disturbing result of meaning nihilism. As Kripke writes, ‘the entire idea of meaning vanishes into thin air’ (1982: 22).

The sceptic’s challenge is notoriously hard to meet. It does no good to observe that our present behaviour is consistent with our past behaviour, as our past behaviour underdetermines the function denoted. Our previous behaviour has been equally consistent with quaddition and addition.<sup>33</sup> If we say that all along we have been following a general rule of addition, rather than relying on a finite stock of previous uses, the sceptic can respond that whatever words or symbols we use in order to express our commitment to the ‘+’ rule, rather than the ‘⊕’ rule, are equally open to deviant interpretations. For example, if the ‘+’ rule included the word ‘count’, might not ‘count’ really have meant ‘quount’?<sup>34</sup> The general problem is that a rule does not determine its own use. It has no intrinsic normative force such that when faced with (the symbols representing) a rule we are obliged to apply it one way rather than another.

For this reason we might be inclined to suppose that the normative force of a rule derives from reasons external to it. Although, we might admit, there is nothing intrinsic to a rule forbidding deviant or non-standard uses, might there not be

---

<sup>33</sup> There are evident parallels here with Hume’s problem of induction and Nelson Goodman’s (1955) ‘new riddle of induction’, the ‘grue’ paradox. See further Kripke (1982: 58–9; 107–8).

<sup>34</sup> See Kripke (1982: 16).

some outside guidance, some ‘infinitely long rails’ which ‘correspond to the unlimited application of a rule’ (*PI* §218)? Wittgenstein denies this:

How can we know how he is to continue a pattern by himself – whatever instruction you give him? – Well, how do I know? – If that means “Have I reasons?” the answer is: my reasons will soon give out. And then I shall act, without reasons. (*PI* §211)

An appeal to external guidance fails because justifications must end somewhere and we are then confronted with the same difficulties. Moreover, the phenomenology of rule-following fails to support the idea that one looks outside the rule itself to determine its correct application: ‘we look to the rule for instruction and do something, without appealing to anything else for guidance’ (*PI* §228). In any case, it would be hopeless to seek guidance from other rules or reasons, since these too are lifeless signs open to many and varied interpretations.<sup>35</sup> Unless we strike bedrock, we are pulled into a hopeless sceptical regress.

A different kind of response to the sceptic, and the one which has attracted most subsequent critical interest, involves an appeal to dispositions.<sup>36</sup> Although there was no past time at which we actually added 68 and 57, had we been asked to do so we would have answered 125. The fact that we were previously disposed to add not quadd is the fact which answers the sceptical challenge. As Wittgenstein has one of his interlocutors say, ‘If I had have been asked what number should be written after 1000, I should have replied “1002”’ (*PI* §187). The primary difficulty with this line of response is that it seems poorly equipped to handle the normativity of meaning.<sup>37</sup> Kripke states the problem:

---

<sup>35</sup> See further McGinn 1984: 18-24.

<sup>36</sup> Most commentators discuss dispositions at some point. Kripke’s discussion comes at pp. 22–37. In defence of a dispositionalist account see Forbes (1984). See also Boghossian ([1989] 2002: 164–71), Martin and Heil (1998) and Kusch (2006: Ch. 3).

<sup>37</sup> A further issue raised by Kripke is the disparity between the infinite number of sums and the finite number of my dispositions, casting doubt on whether my dispositions could suffice to determine my meaning addition rather than quaddition (1982: 26–7). However, as Blackburn ([1984] 2002: 35) argues, it is far from clear whether dispositions are finite or infinite: ‘The brittleness of a glass is a respectable dispositional property. But there is an infinite number of times and places and strikings and surfaces on which it could be displayed.’ Moreover, it seems that the issue of finitude applies not only to the individual, but also to the community at large (see Blackburn [1984] 2002: 37). This threatens Kripke’s own ‘sceptical solution’. See also Forbes ([1984] 2002: 24–6).

A candidate for what constitutes my meaning one function, rather than another, by a given function sign, ought to be such that, whatever in fact I (am disposed to) do, there is a unique thing that I should do. (1982: 24)

For a sign to be meaningful, it must be governed by certain normative standards, say rules, which determine correct or incorrect usage.<sup>38</sup> However, as McDowell sums up the problem, ‘a disposition is not something to which its exercises are faithful’ ([1984] 2002: 50). One way to put the matter is to observe that in order to be governed by a rule it is not sufficient for one to be disposed to conform to it. Rather, one must follow the rule, where following a rule is a matter of conforming to it for the right reasons or in the right way. Of course, providing a satisfactory account of what constitutes the ‘right’ reasons or the ‘right way’ is no easy matter. Roughly speaking, however, we might say that it is a matter of acting in accordance with a rule because of the rule, for the sake of the rule, or, as Kant would say, out of respect for the rule. Thus, following a rule is an intentional act. This is of course the point Kant relies upon in order to maintain that the only thing that is good without qualification is the good will; a good nature is merely beautiful.<sup>39</sup>

A further difficulty for a dispositionalist account is that one can be systematically mistaken in the way in which one is disposed to follow a rule.<sup>40</sup> Consider, for example, Aristotle’s arguments concerning happiness (*eudaimonia*). For Aristotle, most people are systematically mistaken with respect to the meaning of happiness:

The many, the most vulgar, would seem to conceive the good and happiness as pleasure, and hence they also like the life of gratification. In this they appear completely slavish, since the life they decide on is a life for grazing animals. *NE* 1095b18-21

---

<sup>38</sup> The standard view is that the normativity involved is *sui generis* semantic normativity. For a dissenting voice see Wikforss (2001). An alternative view is that language use is governed by pragmatic norms; ‘if you want to get on in society, pass exams, engage in commerce etc. then speak in the following way’. Kripke’s own ‘sceptical solution’ seems to be of this sort. I touch on this point again in Ch. 4.

<sup>39</sup> This may also be one of the intuitions underpinning Searle’s ‘Chinese Room’ thought experiment. Searle’s Chinese Room operator conforms to his instructions, but does not follow them in the requisite sense and, so, fails to grasp of the meaning or significance of their output.

<sup>40</sup> See Kripke (1982: 28–30).

In fact, happiness turns out to be, primarily, if not exclusively, contemplative activity (*theoria*) (NE 1178b29-30). This is so despite the majority's dispositions to use 'happiness' differently. Aristotle has arguments to show that the masses ought not to have the dispositions they do in fact have. They misunderstand the meaning of happiness. Whatever one thinks of Aristotle's arguments, the contrary dispositions of the masses do not render them incomprehensible. There is a logical gap between how people are disposed to use 'happiness' and how people ought to use 'happiness' which it makes sense for Aristotle to attempt to exploit. It makes sense to suppose that someone could be brought to realise that they have been using 'happiness' incorrectly and, hence, change their dispositions accordingly. This is possible only on the assumption that there is some external standard of correctness independent of the disposition they happen to have.

In light of the difficulties facing the various candidates for facts of the matter such that one means plus rather than quus, the sceptic's conclusion is meaning nihilism: there is no fact of the matter in virtue of which sentences such as "the student meant addition by '+'" are true. It emerged from our look at the topic of dispositions that at least one of the missing facts is whatever grounds the normative dimension of meaning. Although we cannot conclusively rule out alternative explanations (perhaps we have missed a non-normative fact which would satisfy the sceptic), it seems likely that the problem of missing normative facts underpins the sceptical paradox. This appears to be Kripke's view:

All [attempts to meet the sceptical paradox] fail to give a candidate for a fact as to what I meant that would show that only '125', not '5', is the answer I *ought* to give. (1982: 11, emphasis added).

This, then, is the sceptical paradox. When I respond in one way rather than another to such a problem as '68+57', I can have no *justification* for one response rather than another. (1982: 21, emphasis added)

Endorsing the sceptical conclusion is a clearly unattractive option. Kripke calls it 'insane and intolerable' (1982: 60). Following the helpful Humean distinction between 'straight' solution' and 'sceptical' solutions, we can divide responses

into two categories (Kripke 1982: 66).<sup>41</sup> A straight solution attempts to show that the initial scepticism was, on closer inspection, unwarranted. Thus, one might try to show that, contrary to initial appearances, there are some facts of the matter by dint of which one means plus rather than quus. A sceptical solution, by contrast, begins with the concession that the sceptic's case is unanswerable. It goes on to maintain that our normal way of doing things is nevertheless legitimate because it does not stand in need of the justification which the sceptic demanded. For instance, a demand for truth-makers might be watered down to a demand for warranted assertability, or the demand for truth conditions might be abandoned completely in favour of expressivism.<sup>42</sup>

We have then a sketch of the rule-following considerations. Attending to the vast critical literature on the intricacies of the argument, and proposed solutions to the sceptical paradox, would require a thesis length treatment at least. As with Hume and Moore above, however, I am more concerned at this stage with the general structure of the problem than with proffered solutions. In particular, I should like to emphasise certain common features of the three sets of considerations, diagnose a common ailment, and, in doing so, set out the main problematic of the thesis.

#### *1.4 Naturalism and Classical Realism*

By way of an opening, it will be helpful to consider Michael Dummett's (1959) claim, endorsed by Kripke (1982: 73), that Wittgenstein's rule-following considerations constitute an attack on 'classical realism'.<sup>43</sup> Dummett characterises classical realism as follows:

It is certainly part of the meaning of the word "true" that if a statement is true, there must be something in virtue of which it is true. "There is something in virtue of which it is true" means: there is something such that if we knew of it we should regard it as a criterion (or at least as a ground) for asserting the statement. The essence of realism is this: for any statement

---

<sup>41</sup> Alternative labels are 'direct' and 'diagnostic' anti-scepticism respectively. See Kusch (2006: 16).

<sup>42</sup> See Kripke (1982: 77–8). See also Blackburn ([1984] 2002: 30–1) for more on this strategy of 'lowering the truth condition', and its relationship to Hume's account of causation.

<sup>43</sup> An equally attractive alternative label is 'mirroring realism'. Heal (1989), borrowing the term from Rorty, develops the implications of mirroring realism at some length in the context of both the rule-following considerations and Quine's theses concerning the indeterminacy of translation.

which has a definite sense, there must be something in virtue of which either it or its negation is true. (1959: 335–6)

Mindful that it may be impossible to characterise classical realism in a manner satisfactory to all parties to the debate, let us try to flesh this out. In particular, what, on the classical realist view, is involved in asserting some normative proposition  $p$ , say, ‘wilful murder is vicious’? We might approach the question by noticing that for realists  $p$  is a possible object of knowledge. On anything like the traditional tripartite conception of knowledge, it follows that  $p$  is truth-apt, possibly true and meaningful. Insofar as  $p$  is a declarative sentence in the business of stating how things stand in the world, it is truth-apt. That normative propositions are truth-apt is independently plausible, though controversial, in light of the well-known Frege-Geach considerations concerning the embedding of normative propositions in truth-functional contexts. In order for  $p$  to be truth-apt, it must be meaningful. Moreover, if  $p$  is a possible object of knowledge, it follows that the belief that  $p$  demands justification and that this demand can be satisfied under certain conditions. The meaning of  $p$  and its justification turn out to be intimately connected. For, in general, some proposition  $p$  is meaningful in virtue of its truth conditions.<sup>44</sup> The truth conditions of  $p$  are given by its truth-makers. On the classical realist view under discussion, these are mind-independent facts to which  $p$  corresponds if true. Moreover, these mind-independent facts justify our belief that  $p$  and validate our claim to know that  $p$ . As Dummett notes, they serve as ‘a criterion (or at least as a ground) for asserting the statement’ (1959: 336).

Classical realism, thus, involves a commitment to a cluster of natural and intuitively attractive views. One of its central elements is the idea that for a sentence, or the proposition which it expresses, to be meaningful, there must be some fact which makes it so. However, it is precisely this thought which the rule-following considerations challenge. The sceptic’s conclusion was that there is no fact of the matter in virtue of which I mean plus not quus. For this reason Kripke endorses Dummett’s view of the import of the rule-following considerations:

---

<sup>44</sup> See Kripke (1982: 70): ‘The simplest, most basic idea of the *Tractatus* cannot be dismissed: a declarative sentence gets its meaning by virtue of its truth conditions.’

If Wittgenstein is right, we cannot begin to solve it [the sceptical paradox] if we remain in the grip of the natural presupposition that meaningful declarative sentences must purport to correspond to facts. (1982: 78–9)

In this way, Kripke motivates his favoured non-factualist sceptical solution.

Such a retreat from classical realism should not, however, be treated lightly. One concern is that the abandonment of classical realism for some domain suggests a downgrading of the subject matter. For instance, rejecting moral realism might seem to lead to the (morally) disastrous step of rejecting the reality of moral obligation. Perhaps the disaster can be averted. In part, this perceived threat motivates sophisticated non-cognitivist positions such as Simon Blackburn's quasi-realism. In any case, it is a plausible desideratum that an account of normative discourse should be justificatory in the sense that it is consistent with our continued entitlement to employ normative discourse more or less as our current reflective practice suggests.

Do the rule-following considerations force us to reject classical realism in favour of non-factualism? The challenge was to locate the truth-makers for propositions such as 'Jones meant addition by "+"'. Kripke's Wittgenstein argues that the search fails. Dispositionalism was perhaps the most promising approach, but we saw it blocked by the normative dimension of meaning. The difficulty is that we need not only truth-makers, but also what we might call 'value-makers'. By value-makers I have in mind a special category of truth-makers. They are truth-makers insofar as they are facts or states of affairs which ground the truth of normative propositions. However, they also make certain claims on us. Rules are putative examples of value-makers. If we are to follow rules, they must, in some sense, lead. Value-makers have normative force; they are compelling standards.

The present thesis is, in part, an enquiry into the metaphysics of value-makers. It is also an attempt to assess the prospects of such an enquiry governed by naturalistic constraints. Each of the three sets of considerations discussed in this chapter was constrained in this way. Hume is an explicit naturalist and generates

for himself the problem of locating the value-makers which would give content to our normative ideas. Moore develops the problem of identifying natural properties with the appropriate normative force in order to challenge naturalist opponents. Although Kripke's Wittgenstein is a more ambiguous creature, the sceptical paradox gets off the ground only once we couple it with naturalistic sounding assumptions about what sorts of facts we should look for to determine meaning. Although Kripke initially presents the argument in an epistemological guise, he argues that the problem is ultimately metaphysical, allowing us the possibility of meeting the sceptic from a gods-eye perspective.<sup>45</sup> Thus, he tells us 'there are no limitations ... on the facts that may be cited to answer the sceptic' (1982: 14). However, 'no limitations' appears to be an exaggeration. The limitations Kripke has in mind are behaviourist limitations, sometimes associated with Wittgenstein's philosophy of mind. However, naturalistic limitations remain. Kripke's discussion is in fact limited to properties such as dispositions, mental histories, qualitative states and the like. As Jane Heal writes:

Our looking for these kinds of items in answer to the original questions about thought is bound up with the dominance of scientific modes of understanding and our disposition to assume that we can find an account of thinking closely analogous to the kinds of account which science has given us of the constitution and workings of the natural world. (1989: 146)

There remains the unexplored possibility of an appeal to a non-natural property with the appropriate normative force. Perhaps, for example, God guarantees that we mean plus not quus. Perhaps meaning-constituting facts are to be discovered in a Platonic realm of Ideas. Or, perhaps, primitive *sui generis* semantic properties justify us in our meaning practices.<sup>46</sup> Each of these would serve simultaneous duty as truth-makers and value-makers.

---

<sup>45</sup> This view of the argument is endorsed by Boghossian: 'once we have corrected for the distortions induced by the dialogic setting, there ought not to be any residual temptation to think that epistemological considerations are playing a critical role in Kripke's argument.' ([1989] 2002: 151)

<sup>46</sup> Kripke briefly raises the latter possibility, but rejects it as 'desperate' and 'completely mysterious' (1982: 51). See also Boghossian ([1989] 2002: 178–80).



As I have emphasised, the open question remains open, and Kripke's sceptic remains unsatisfied, due to the phenomenon of normative force. Heal comments on this fact in relation to the rule-following considerations as follows:

What Wittgenstein's considerations show is that we cannot make sense of the idea that being confronted with some set of objects or chunk of the world should *force* possession of a certain concept upon a person.' (1989: 146–7, emphasis added)

We can perhaps appeal to chunks of the world in order to provide efficient causal explanations of concept acquisition. This is one sort of force. However, we need a further sort of explanation for the queer phenomenon of normative force. If we are committed to classical realism, we need to locate correspondingly queer properties, namely value-makers. We have reason to think that a non-queer explanation, in terms of dispositions for instance, will not do. Even if our dispositions to employ concepts in particular ways, or the categorical bases of these dispositions, causally determine future use, they do not thereby provide normative explanations. As Paul Boghossian writes:

Even if there were a dispositional predicate that logically covaried with a meaning predicate, the one fact could still not be identified with the other, for they are facts of distinct sorts.' ([1989] 2002: 169).

Recall the earlier quotation from Moore:

It may indeed be true that what we desire to desire is always also good; perhaps, even the converse may be true: but it is very doubtful whether this is the case, and the mere fact that we understand very well what is meant by doubting it, shews clearly that we have two different notions before our minds. (*PE* 16)<sup>47</sup>

The OQA and the rule-following considerations converge at this point. Naturalists must either fully account for normative properties in terms of natural properties or eliminate them from their ontology and abandon realism with respect to them. However, the queer phenomenon of normative force presents an obstacle to naturalistic accounts of normative properties. One way to show ontological incompleteness is by appeal to independent variation. For example, given that two points can coincide in three dimensions and still vary with respect

---

<sup>47</sup> See also *PE* 14: 'Even if it [good] were a natural object, that would not alter the nature of the fallacy nor diminish its importance one whit'.

to location, we know that we cannot provide a complete account of location in terms of three dimensions.<sup>48</sup> If normative properties vary independently of natural properties, then the latter cannot offer a complete account of the former. The OQA and the rule-following considerations, if compelling, show that even were some natural property shown to co-vary with some normative property in terms of extension, this fact would not thereby demonstrate that they co-vary in terms of normative force. For, we can always coherently question whether any empirically given property motivates us appropriately. Similarly, we have seen Hume's argument that the content of our normative ideas cannot be grounded in our experience of the external world, for there is nothing in our experience which resembles the element of approbation or disapprobation. The shared thought is that our experience of the external world is necessarily compatible with motivational indifference.

If we want to retain classical realism with respect to the normative, we need value-makers in order to make sense of normative propositions and sentences. We need to ask 'do value-makers exist, and, if so, how do they fit into our ontological scheme?' Whilst this is the question in full generality, my current ambition is to investigate how the problem becomes manifest within naturalistic constraints. As I shall argue in the following chapter, naturalism is an epistemological or methodological position. It is the adoption of a scientific stance towards traditionally philosophical problems. How is this epistemological stance to be reconciled with the problem of value-makers?

Given the difficulties involved in locating value-makers in the mind-independent external world, one possibility is to relocate them in agents' subjective mental states. For instance, we might turn to instrumentalism, associating value-makers with the satisfaction of agents' desires. We ought to act in such and such a way, or employ concepts in such and such a way, because doing so will enable us to meet our ends. This approach has the advantage of appearing to explain why it is that norms have motivational force. For, it is an apparent truism that we are motivated to bring about what we desire. It is also perhaps the most natural view of the matter for a thoroughgoing naturalist, dovetailing with their available

---

<sup>48</sup> This example is taken from Frank Jackson (1998: 9).

resources. In chapter five I consider the merits of this approach as advocated by Quine, Maddy, Papineau and Kornblith with respect to epistemic normativity. I shall argue, however, that instrumentalists err in dismissing the categorical force of certain norms. I shall also argue in chapter eight that it is a mistake to ground reasons in desires.

A further possibility is to argue that an unduly restrictive naturalistic methodology has caused us to overlook mind-independent value-makers. Such, of course, was Moore's view. It was also a view towards which Wittgenstein, or one of his interlocutors, was tempted:

“But I don't mean that what I do now (in grasping a sense) determines the future use *causally* and as a matter of experience, but that in a *queer* way, the use itself is in some sense present.” (*PI* §195)<sup>49</sup>

Although a brute appeal to intuition of non-natural properties, or to grasping ‘in a flash’ (*PI* §191), does little to lessen the sense of mystery, I think that these answers lie in approximately the right direction. While normative facts are not revealed by the methods of the natural sciences, they are nevertheless real and cognitively available. The appropriate epistemology, I shall argue in chapter nine, depends upon the emotions. But, whilst the emotions can provide the warrant for normative judgments, they do not provide the grounds. Rather, I shall argue that the value-makers of normative propositions are aspects of human practices combined with facts about our nature as rational agents. I shall argue for this positive account in chapters six and seven especially. Given the emphasis on practices, there are certain affinities between my position and Wittgenstein's emphasis on ideas such as forms of life. However, the account I develop draws rather on the Aristotelian and neo-Aristotelian traditions, placing emphasis on the role of the virtues and the teleological structure of human activities.

---

<sup>49</sup> There is an interesting coincidence of vocabulary here between Wittgenstein (as translated by Anscombe) and Mackie's argument from queerness. Bloomfield (2006: 185) also observes this point.

## **Chapter Two** **Naturalism as an Epistemological Stance**

In the previous chapter, I set out a problem for naturalistic accounts of normativity via a discussion of three influential sets of considerations and their interconnections. In doing so I was content to rely on a relatively imprecise, intuitive sense of what naturalism is. Given the differences and ambiguities involved in the three underlying conceptions, greater definition would have tended to obscure the central issue in any case. In order to begin the process of understanding why naturalism struggles to accommodate the normative, however, it will be helpful to consider the nature of naturalism in far greater detail. The task is independently important given the popularity of naturalism in contemporary analytic philosophy.<sup>1</sup> Indeed, the naturalistic turn has taken such a hold that challenging a theory's naturalistic credentials is widely regarded as a legitimate form of criticism. This might have something to do with the way in which the term 'naturalism' has come to take on positive connotations. Naturalism is associated with healthy respect for the explanatory, predictive and technological successes of the natural sciences and due modesty regarding 'armchair' philosophy. The progress of the natural sciences, considered in juxtaposition to the apparent lack of progress in traditional philosophy, cautions that favouring philosophical arguments over scientific research is likely to be unwise.<sup>2</sup> Insofar as this caution remains unheeded, non-naturalism appears to be a headstrong attitude. It is thus unsurprising that philosophers are frequently hesitant to identify themselves as 'non-naturalists'. Accordingly, however,

---

<sup>1</sup> In a recent straw poll of 931 working academic philosophers, 49.8% of respondents leaned towards or accepted naturalism, as opposed to 25.8% of respondents who leant towards or accepted non-naturalism. The remaining 24.2% of respondents selected 'other', presumably on the grounds that the term 'naturalism' is unclear. The results, along with meta-analysis, can be found online at <http://philpapers.org/surveys/results.pl>

<sup>2</sup> David Lewis puts the point rather more provocatively: 'I am moved to laughter by the thought of how presumptuous it would be to reject mathematics for philosophical reasons. How would you like the job of telling the mathematicians that they must change their ways, and abjure countless errors, now that philosophy has discovered that there are no classes? Can you tell them, with a straight face, to follow philosophical argument wherever it might lead? If they challenge your credentials, will you boast of philosophy's other great discoveries: that motion is impossible, that a Being than which no greater can be conceived cannot be conceived not to exist, that it is unthinkable that anything exists outside the mind, that time is unreal, that no theory has ever been made at all probable by evidence (but on the other hand that an empirically ideal theory cannot possibly be false), that it is a wide-open scientific question whether anyone has ever believed anything, and so on, and on, ad nauseam?' (1991: 59).

philosophers are inclined to view naturalism through the prism of their independently favoured commitments. While it sometimes seems as if there are as many naturalisms as there are self-proclaimed naturalists, there is at least a reasonably clear division between ontological and epistemological naturalisms. In the first section of this chapter I shall argue that insofar as naturalism is a novel development in philosophy, it is fundamentally an epistemological thesis, without particular ontological commitments. In the second section I shall consider the charge that epistemological naturalism is self-defeating and defend the desirability of a pragmatic criterion for adjudicating between naturalistic and non-naturalistic approaches to philosophical problems.

### *2.1 Epistemological and Ontological Naturalism*

When naturalism is advanced as a positive thesis, two views are frequently distinguished, namely epistemological naturalism (EN) and ontological naturalism (ON).<sup>3</sup> EN, or the pejorative ‘scientism’, is the view that knowledge can only be acquired through the natural sciences.<sup>4</sup> It is the view that ‘there is only one way of knowing: the empirical way that is the basis of science (whatever that may be)’ (Devitt 1998: 45). Or, as Bertrand Russell put it, ‘what science cannot discover, mankind cannot know’ (1935: 243).<sup>5</sup> Certainly we can come to have true beliefs on the basis of metaphysical argument, astrology or being struck on the head by a falling object, but these true beliefs are not justified and, therefore, not knowledge. The second thesis, ON, holds that reality is exhausted by some combination of natural objects, properties, relations and events. For example, David Armstrong advances the view that ‘the world, the

---

<sup>3</sup> Axiological scientism is sometimes further distinguished, as the view that science is the most valuable part of learning and culture (see, for example, Stenmark 2001: 11–13).

<sup>4</sup> A stronger claim would be that we are rationally entitled to believe only what is scientifically knowable. The conditions for knowledge and rational belief come apart in instances where we have good reason to believe what is in fact false. For example, we might think that people were once rationally entitled to believe that the earth was flat or that space was globally Euclidean.

<sup>5</sup> Carnap advances the yet stronger thesis that ‘there is no question whose answer is in principle unattainable by science’ (1967: 290; emphasis removed). Idealism aside, this might seem to be a rather naïve expression of optimism. For, what reason do we have to suppose that humans are capable of knowing all there is to know? In Carnap’s case the claim is grounded in his meaning-empiricism, according to which the class of intelligible statements is exhausted by those statements which can in principle be verified by experience. As the class of meaningful questions is exhausted by those which can be answered by intelligible statements, all meaningful questions can be answered by science.

totality of entities, is nothing more than the spacetime system' (1997: 5).<sup>6</sup> Similarly, Jaegwon Kim tells us that 'the physical world of spacetime is the whole world' (2005: 157). Although ON is frequently combined with the claim that all natural properties are reducible to physical properties, ON and physicalism are in fact distinct. For instance, whilst Armstrong is a physicalist, he distinguishes his naturalism from his physicalism on the grounds that naturalism is not committed to the physical world being ontologically fundamental. For Armstrong, the view that spacetime is ultimately analysable in terms of non-physical entities would still count as naturalistic. Nevertheless, Armstrong's naturalism is a substantive thesis concerning what exists and, thus, an ontological thesis.

It might be thought all naturalists have some definite ontological commitments insofar as they uniformly reject the existence of the supernatural; spooky entities such as Cartesian souls, ghosts and spirits. This is helpful only to the extent that nature can be distinguished from supernature and I shall look at some of the difficulties involved in drawing the distinction below.<sup>7</sup> However, it is not clear in any case that a disavowal of the supernatural is sufficient for naturalism. Wherever the line between nature and supernature is finally drawn, we might expect God and transcendent universals to fall on the wrong side for the naturalist. Let us imagine, for the sake of simplicity, that the putatively supernatural is exhausted by these two categories and consider a philosopher who denies the existence of entities falling under both. Are we thereby entitled to categorise her as a naturalist? It seems that we are not yet in a position to know. The question we need to answer is not which entities appear in her ontology, but on what grounds they appear. It may be that she is persuaded by, say, the problem of evil and the third man argument. In this case it would be a mistake, I suggest, to conclude that she is a naturalist simply because she has no ontological commitments to the supernatural. Unlike naturalists, she does not reject God and Platonic Forms on the basis of a principled opposition to the supernatural.

---

<sup>6</sup> See also Armstrong (1980: 149).

<sup>7</sup> As Berkeley observed, 'nature' is apt to become 'an empty sound, without any intelligible meaning annexed to it' (*Principles*, paragraph 150). Hume says of the term 'nature', that 'there is none more ambiguous and equivocal' (T III, 1, II, 7).

The difference emerges most clearly when we consider a shift in her ontological commitments. Let us say that she currently holds mathematical objects to be naturalistically explicable. Perhaps she holds an error theory of mathematics or thinks that mathematical discourse can be reduced to talk of sets constituted by physical objects.<sup>8</sup> On further reflection, however, she reaches the conclusion that a particular form of mathematical Platonism is unavoidable. She now comes to think that an appeal to transcendent universals is the only way to preserve the *a priori* nature of mathematical knowledge, or to explain the necessity of mathematical truth. As she had no previous ontological commitments to supernatural entities, must she avoid Platonism on pain of inconsistency? Given that her previously naturalistic ontology was not based on an over-riding naturalistic principle, there is no reason to think so. There is nothing to prevent her transition from a natural to a supernatural ontology. If, however, she had been a *bona fide* naturalist, with naturalism understood in terms of an opposition to a non-supernatural ontology, she would have had no such leeway. She would then have been forced either to eliminate mathematical properties from her ontology or to redraw the boundary between nature and supernature.

According to the above line of argument, then, a non-supernatural ontology is not sufficient for naturalism. In due course I shall also argue that it is not necessary.<sup>9</sup> First, however, I would like to consider how we might try to distinguish the natural from the non-natural, or the physical from the non-physical. As mentioned above, this is no easy task. For instance, it is unhelpful to say that something is physical if and only if it is extended in space, when we have an inadequate idea of how to characterise space independently of physical extension.<sup>10</sup> It would also be implausible to claim that a property is natural if and only if it figures in the theories of natural scientists, or that the physical is

---

<sup>8</sup> See Maddy (1990) for a defence of this view.

<sup>9</sup> Hence, I do not share van Inwagen's sense that 'we should certainly be very puzzled if someone said, "I'm not a naturalist, but I'm a physicalist" – or "I'm not a physicalist, but I'm a naturalist"' (2006: 78).

<sup>10</sup> A further problem with the proposal is that would classify any irreducibly mental properties with spatial locations as physical. For further discussion of defining the physical in terms of spatial extension see Crane (1993) and Pettit (1993) and Moser and Yandell (2000: 4). Markosian (2000) offers a defence of the proposal.

coextensive with the set of entities figuring in the theories of respected contemporary physicists. The history of science gives us no good reason to suppose that early twenty-first century physicists have identified all and only the fundamental constituents of reality, and plenty of reasons to suppose otherwise given the changing subject matter of the discipline and the present lack of ontological consensus.<sup>11</sup> Furthermore, this view would have the unpalatable consequence that any future extensions of physics would not describe the physical.<sup>12</sup> Quite generally, moreover, it seems wrong-headed to think that ontological facts are determined by physicists' theories of reality, rather than by reality itself. At least, few naturalists would be content to embrace idealism. All of this glosses over the considerable difficulties concerning how a scientist's ontological commitments are to be read off from her theory.<sup>13</sup>

In light of such considerations, it is sometimes suggested that we should place our ontological commitments on the side of a hypothetically completed physics of the future.<sup>14</sup> For instance, one might claim that a predicate, such as 'is a quark', succeeds in picking out some aspect of reality just in case an ideal physics would rely on that predicate in its foundations. Setting aside the question of whether we have good reason to believe that physics in anything like

---

<sup>11</sup> See Ritchie (2008: 95-6) for more on naturalism and the so-called 'pessimistic meta-induction'. Attempting to give some content to the term 'physicalism', Kim has claimed that 'the broad basic features of the world as described by modern physics, what is intelligible and of interest to those of us who are not science specialists, has been relatively stable through the flux of changing physical theories' (2005: 149). I am not similarly sanguine concerning the stability of theoretical physics. Although I do not know exactly what Kim means by 'modern', it is not difficult to think of substantial revisions within fundamental physics over most timeframes. What seems especially odd, however, is Kim's appeal to the picture of physics held by people who are not specialists. If the appeal is to folk physics, then this stands apart from the cutting edge of the scientific enterprise. Indeed, I would imagine that the folk conception has been more or less impervious to theoretical advances since around the time of Bohr's model of the atom. If, however, physicalism is supposed to derive support from the sciences, as Kim thinks it does, then it is hardly satisfactory to draw that support from ideas independent of the 'flux of changing physical theories', which is to say independent of our best current theories. Ladyman and Ross have recently lampooned elements of contemporary analytic metaphysics as the 'philosophy of A-level chemistry' (2007: 24). Regardless of the justness of their critique, it must be recognised that there is a pitfall to be avoided. If one's philosophical claims depend on premises drawn from the physical sciences, it does not do to proclaim them independent of any possible scientific advance.

<sup>12</sup> As Bas van Fraassen writes: 'Whenever philosophers take some general feature of physics and use it to identify what is material, what happens? Physics soon goes on to describe things that lack that feature and are altogether different. When that happens, does materialism bite the dust? Surely not!' (2002: 53). See also Crane and Mellor (1990: 188).

<sup>13</sup> See Raley (2005) for more on this issue.

<sup>14</sup> See, for example, Smart (1963: 651), Petit (1992: 247) and Papineau (1993: 2).



its current form is in principle completable, it seems that the grounds on which we might be inclined to appeal to an ideal physics are once more epistemological rather than ontological. As we have no guarantees concerning which entities a hypothetically completed physics of the future may contain, the idea seems to be rooted in the belief that the methods of the sciences will ultimately reveal the constituents of reality, whatever they may turn out to be. For the present, however, we cannot be certain that the current subject properties of physics will turn out to be physical, for we cannot yet know that properties such as spin, mass and charge will feature in an ideal physics.<sup>15</sup> If inclusion in a hypothetically completed physics of the future is advanced as the criterion for a substance or property being physical, we currently find ourselves in an uncomfortable position from which we can only speculate on the question of what is, and what is not physical.<sup>16</sup> We cannot be sure, moreover, that an ideal physics will exclude paradigmatically supernatural entities. For example, some philosophers, impressed by the argument from design, have argued that the existence of God can finally be confirmed by the natural sciences. Thus, Richard Swinburne claims that the case for theism has the same structure as the ‘cumulative case for any unobservable entity, such as a quark or a neutrino’ (1983: 386).<sup>17</sup> Now, it may of course turn out that Swinburne is wrong to suppose that science will ultimately support theism. Perhaps physics will finally explain everything in terms of atoms and the void or strings vibrating in eleven dimensions. Nevertheless, the current state of scientific knowledge does not rule out the possible existence of any particular entity. It cannot do, for scientific knowledge is contingent and corrigible. Quine is particularly alert to this point:

Even telepathy and clairvoyance are scientific options, however moribund. It would take some extraordinary evidence to enliven them, but if that were to happen, then empiricism itself – the crowning norm ... of naturalised epistemology would go by the board. For remember that that norm, and

---

<sup>15</sup> Crane and Mellor (1990: 188) advance a similar argument.

<sup>16</sup> On a similarly gloomy note Hilary Putnam observes that ‘the idea that science leaves no room for an independent philosophical enterprise has reached the point at which leading practitioners sometimes suggest that all that is left for philosophy is to try to anticipate what the presumed scientific solutions to all metaphysical problems will eventually look like’ (1992: preface, x). This leaves traditional philosophy on a highly uncertain footing. Philosophers are not especially well equipped to speculate on such matters, by method, by training and frequently by inclination.

<sup>17</sup> See Swinburne (1991: Ch. 8) for an influential presentation of this argument.

naturalised epistemology itself, are integral to science, and science is fallible and corrigible. (1992: 20)

If I saw indirect explanatory benefit in positing sensibilia, possibilia, spirits, a Creator, I would joyfully accord them scientific status too, on a par with such avowedly scientific posits as quarks and black holes. (1995: 252)

For Quine, this comes as a consequence of confirmational holism and the resulting impossibility of clearly defined *a priori* knowledge. However, it also comes more simply from a commitment to EN.

Notoriously, Quine included *abstracta* such as sets in his ontology because he considered them indispensable to our best scientific theories.<sup>18</sup> One might simply discount Quine as a naturalist on this basis, despite his prominence in the tradition. Jerold Katz takes this line:

Naturalists cannot countenance the existence of universals and abstract objects, whose existence is independent of natural objects, without contradicting their monism, which denies that anything exists over and above natural objects. (1990: 238)

Of course, ‘naturalism’ is a term of art and naturalists are free to characterise it as they see fit. Naturalism can be characterised as physicalist monism, perhaps combined with something like a deep-seated respect for the results of the natural sciences. Still, there is no guarantee that the results of the natural sciences do or will cohere with physicalism.<sup>19</sup> In such cases as there is or could be disagreement, it is pertinent to ask where a naturalist’s commitments would lie and upon what basis. If the naturalist’s commitment to physicalism depends upon its alleged scientific support, then, that support being removed, he ought to

---

<sup>18</sup> See Colyvan (2001) for extended discussion of the Quine-Putnam arguments for the indispensability of mathematical objects to scientific discourse.

<sup>19</sup> Perhaps one could avoid the appearance of such disagreement between ON and science by declaring by fiat that all scientific commitment is to natural properties, because the subject properties of the sciences are limited to natural properties. If this is combined with EN, the view that all genuine knowledge is scientific knowledge, then we can only have knowledge of natural properties. If we should only commit to the existence of those properties which we could in principle have knowledge of, then a commitment to ON follows from EN. However, this is unsatisfactory for at least three reasons. First, it is arbitrary to restrict science to natural properties. Second, we do not have a firm grasp of the distinction between natural and non-natural properties. Third, it follows at best that we cannot have knowledge of non-natural properties, not that they do not exist. For, science by this definition, remains mute on the subject of non-natural properties. Hence, we do not know that non-natural properties do not exist; it is rather that we cannot know that they do exist.

abandon his commitment. If he does not do so, it must be because he has further, non-scientific reasons. That is, most plausibly, he needs a metaphysical argument. Moreover, he must be willing to assign greater weight to the metaphysical argument than to the scientific evidence in this particular case, or see it as tipping the balance against mixed evidence. If this position merits the label ‘naturalism’, however, it is not significantly different from philosophical materialism in the grand old tradition of Democritus and Hobbes. It is true that a naturalist of this stripe will pay especially careful attention to scientific evidence. However, it is only sensible for any philosopher to consider empirical evidence for or against his theses where relevant. Philosophical disagreements do not focus on whether to consider relevant empirical evidence, but on whether the evidence is indeed relevant and how much significance to assign it. By contrast, the form of naturalism endorsed by philosophers such as Quine, the form of naturalism that has given rise to the naturalistic turn and a professed abandonment of first philosophy, is far more radical. The principled epistemological naturalist shows no potential willingness to allow philosophical arguments to trump scientific evidence, for he does not recognise the validity of non-scientific methods. He will follow science, even if science leads him to accept the existence of the supernatural, and he will certainly not lay claim to his title on the basis of a metaphysical defence of physicalism.<sup>20</sup> As van Fraassen writes:

If you press a materialist, you quickly find the most important constraint on the meaning of the thesis. That constraint is simply that it should be compatible with science, whatever science comes up with. (2002: 55–6)

Compatibility, however, is too weak. Assuming that science accurately describes the world within its domain of competence, everyone should intend his or her descriptions of reality to be compatible with science. This is just as if a scientist should hope that his theories are compatible with our best philosophical accounts of the world – if, that is, he grants philosophy jurisdiction over a realm of facts

---

<sup>20</sup> This does not, of course, prevent him from arguing that our best available scientific models suggest that the world is entirely physical and, so, to adopt physicalism as a working hypothesis. In fact, it may turn out that a naturalist, that is an epistemological naturalist, should also endorse physicalism; if, that is, our best scientific models suggest that physicalism is true. What I am denying is that it is in virtue of a philosopher’s endorsement of physicalism that she is to be counted as a naturalist. Rather, if there is any connection between physicalism and naturalism, it is that the former offers some defeasible reason to endorse the latter.

which scientific practice cannot describe. However, the naturalist denies any such philosophical jurisdiction. His quest is thus for scientific dominion, not alliance or mere compatibility.

I have argued, then, that either naturalism is best characterised in epistemological terms, or otherwise that the self-professed naturalist is in fact a philosophical materialist in the grand old tradition. It might be objected, however, that this provides the wrong order of explanation, that the truth of ON is the reason to endorse EN. One argument to this effect starts by noting that the natural sciences exclusively study natural properties. Given ON, there is nothing else to study. If non-natural properties existed, they would lie outside the natural sciences' domain of competence in any case. Let us imagine, however, that ON is false, and, hence, that non-natural properties exist. In this case, we could presumably acquire knowledge concerning at least some of these non-natural properties. At least we could know *that* some non-natural property exists, for its existence would constitute the grounds for rejecting ON. However, if ON were false and we could have knowledge concerning non-natural properties, then EN would also have to be false insofar as it is the view that all knowledge is scientific knowledge. Our knowledge concerning non-natural properties would depend on non-scientific methods. Therefore, if ON is false EN must be false. The truth of EN depends upon the truth of ON.

Mikael Stenmark defends the priority of ON along similar lines

Ontological scientism entails epistemic scientism because we could not know anything about what does not exist. We cannot know anything about a reality to which science does not have access, because there is simply no such reality. (2001: 8)

Aside from the now familiar difficulty of finding a principled distinction between natural and non-natural properties, the problem with this line of argument is that it is viciously circular. If only natural properties exist, and if the only way to acquire knowledge of natural properties is through the methods of the natural sciences, then ON does indeed entail EN. However, the question of whether the natural sciences *and only* the natural sciences can deliver knowledge of natural properties is open to dispute. One might agree that only natural properties exist,

but also maintain that we can gain knowledge of natural properties through non-scientific methods such as traditional philosophical theorising. In later chapters, I seek to defend just such an account of the normative insofar as I make no appeal to obviously non-natural properties. Yet, whilst I shall draw on empirical evidence at certain stages, most of my arguments will be distinctively philosophical. Thus, my account of the normative will serve as a concrete example of the independence of EN from ON. The example is slightly complicated by the fact that I am not, as it happens, willing to identify myself as an ontological naturalist. I do not restrict my normative ontology to natural properties on principled naturalistic grounds. Indeed, as I discuss in chapter six, I have a certain degree of sympathy for the view that my preferred account of normativity demands a larger supernatural framework. Still, the bulk of my discussion is consistent with ON. As a naturalist could endorse my ontology, but not my methodology, EN cannot follow from ON. The argument in favour of the priority of ON would succeed if one had a prior commitment to EN, but that it precisely the point at issue. Thus, the argument poses no threat to our characterisation of naturalism in epistemological terms.

## *2.2 Is Epistemological Naturalism Self-Defeating?*

I have argued, then, that naturalism is best understood as an epistemological position. Attempting to characterise ON, we are left either with a thesis devoid of definite ontological content or with a form of traditional materialism. Ontological considerations are neither necessary nor sufficient for naturalism. EN however faces an immediate and pressing difficulty, for it is apparently a philosophical thesis, not a scientific thesis. As Stenmark writes:

How do you set up a scientific experiment to demonstrate that science or a particular scientific method gives an exhaustive account of reality? I cannot see how this can be done in a non-question begging way. What we want to know is whether science sets the limits for reality. The problem is that since we can only obtain knowledge about reality by means of scientific methods ... we must use those very methods whose scope is in question to determine the scope of these very same methods. (2001: 23)

Science is an inappropriate tribunal to pronounce on its own global applicability. EN is a philosophical thesis. As EN denies the authority of any extra-scientific

tribunal, it cannot be justified by its own lights.<sup>21</sup> This kind of difficulty is familiar to restrictive methodological proposals. Famously, the logical positivists' verification principle is neither empirically verifiable nor a tautology. The verification principle is a descendant of Hume's Fork, which is neither a relation of ideas nor a matter of fact. Relativism faces *peritrope* arguments.

There is a danger of reaching an unproductive standoff. Naturalists are able to accept that EN cannot be justified by distinctively philosophical methods. They claim that distinctively philosophical methods cannot generate knowledge in any case. Hence, *nothing* can be justified on distinctively philosophical grounds, and it is no criticism that EN cannot be thus justified. Non-naturalist philosophers agree that EN cannot be justified on philosophical grounds, but see this as damaging. For they argue that EN is a philosophical thesis, and if it cannot be justified on philosophical grounds, it cannot be justified at all. Both sides of the debate might thus be tempted to disengage. The naturalist offers the non-naturalist no reason to accept her philosophical thesis. By the naturalist's lights, the philosopher's methods, and hence her criticisms, are unjustified and so do not call for rational debate.

It seems that the only way to make sense of naturalism from a philosophical perspective is to suppose that it has some special status such that it does not stand in need of independent justification. Perhaps EN is intended to show us the way and then be kicked away like Wittgenstein's ladder, or to re-orientate our thinking towards methods which can then be justified by their fruits. Hilary Kornblith writes in this vein:

What does have priority over both metaphysics and epistemology, from a naturalistic perspective, is successful scientific theory, and not because there is some a priori reason to trust science over philosophy, but rather because there is a body of scientific theory which has proven its value in prediction, explanation, and technological application. This gives scientific work a kind of grounding which no philosophical theory has thus far enjoyed. (1994: 49)

---

<sup>21</sup> Almeder (1998: 64–74) develops this point at considerable length, considering it decisive against global epistemological naturalism.

The natural sciences, naturalists justly point out, have been enormously successful at ameliorating our epistemic position. Even if the demand for external justification is impossible to meet, an advocate of EN can argue that he is under no obligation to do so.<sup>22</sup> The demand is merely symptomatic of a stubborn refusal to abandon the project of first philosophy, a research programme now superseded by the natural sciences. As Quine writes:

If the epistemologist's goal is validation of the grounds of empirical science, he defeats his purpose by using psychology or other empirical science in the validation. However, such scruples against circularity have little point once we have stopped dreaming of deducing science from observation. (1969a: 75-6)

It might be objected that the naturalistic proposal to exclude distinctively philosophical methods is unduly dogmatic. However, traditional philosophy also places limits upon acceptable methods of enquiry. Enquiry must begin from somewhere, and the methods of enquiry must be limited in some fashion. All enquiries necessarily involve a prior disposition to respect some sources of evidence and not others. For example, philosophers do not accept bare appeals to religious experience as evidence. Perhaps, then, we can understand the naturalist's puzzlement when faced with a demand for philosophical justification in the same way that we can understand a philosopher's puzzlement if confronted with a demand that their rejection of divine revelation be justified by divine revelation. On the word of which Deity, one might ask, do philosophers base their resistance to religious experience? On the basis of which metaphysical argument, the naturalist is asked, do you reject metaphysics?

---

<sup>22</sup> Another possible response to the circularity objection claims that as traditional philosophy has had little success in providing a justificatory framework for scientific enterprise, the naturalist is not, as a matter of fact, in a worse position. Thus, Ronald Giere writes: 'The enterprise of trying to justify science without appeal to any even minimally scientific premises has been going on without conspicuous success for three hundred years. One begins to suspect the lack of success is due to the impossibility of the task. Perhaps there just is no place totally outside science from which to justify science. At the very least one might conclude that the task is not going to be accomplished any time soon by ordinary mortals. I am willing to take this as sufficient grounds for an ordinary mortal to try something else.' (1987: 148) However, there are at least three reasons not to be persuaded by this argument. First, it is unclear that the mainstream of traditional philosophy of science is in such poor shape as Giere contends. Hotly disputed questions in epistemology, as elsewhere in philosophy, tend to attract the most attention and obscure domains of broad consensus. Second, if anything Giere's argument tends to support skepticism rather than naturalism. Third, if the circularity objection works, it is decisive against naturalism. If Giere's response works, it only shows that the traditional approach has not yet succeeded. Thus, there is an asymmetry between the initial objection and the *tu quoque* response.

It might seem, however, that there is an important difference between the two cases. EN is a philosophical thesis and therefore demands philosophical justification. By contrast, a prohibition on religious experience as a source of evidence is not obviously a religious thesis, demanding religious justification. Indeed, it appears to be another philosophical thesis, concerning religion this time, rather than science. From a theistic perspective, however, it is not at all clear that we are dealing with a non-religious thesis. For, a theist might argue, the cogency of our commitment to various methodological research programmes can only be assessed and justified from within a theistic framework. It may be, for example, that the omniscience and omnibenevolence of God, revealed directly through religious experience, is the only way in which to justify appeals to religious experience. A non-theistic philosopher will certainly challenge the epistemological grounds for this assertion. The theist, however, will argue that he lacks the resources to do so. This mirrors the case of the naturalist marshalling scientific evidence for naturalism, without, the non-naturalist philosopher will say, the appropriate metaphysical framework. We are drawn once more into a dialectical impasse. Regarding EN as the statement of a research programme, to be assessed according to pragmatic criteria, might afford us greater hope of progress.

As we have said, explanations and methods of justification must end somewhere. This does not mean, however, that it is a purely arbitrary matter where our spade turns. Rather, we can pause, take stock and consider whether our methods are conducive to the end of enquiry, say knowledge.<sup>23</sup> If we come to see that our methodological principles are not reliably conducive to knowledge acquisition in some domain, the rational course of action is to consider making changes to our set of methodological principles.<sup>24</sup> This is not to suggest the possibility of

---

<sup>23</sup> In chapter seven, I focus on the end of enquiry in some detail.

<sup>24</sup> Katz makes a similar point. He begins by distinguishing three levels of theory, namely philosophical theories, metatheories and object theories. Following his terminology, we might say that naturalism is a philosophical theory, which translates into a metatheory placing various constraints upon the special sciences, the contents of which are object-theories. He then argues as follows: 'If the constraints of a metatheory prove too strong for the construction of adequate object-theories, the metatheory loses creditability, and, as a consequence, the philosophical theory that sanctions the metatheory loses creditability, too. For instance, when psychological theories constructed in accord with behaviourist metatheory turned out not to handle facts



distancing ourselves from all of our methodological principles and considering them from neutral ground. The evaluation of methodologies is itself a form of enquiry and as such involves prior dispositions to recognise some forms of evidence and not others. We cannot, for instance, distance ourselves from our entire stock of logical principles in order to submit them to independent, non-logical scrutiny. We can, however, learn from the successes and failures of others who are already engaged in alternative research programmes.<sup>25</sup>

The natural sciences have proven successful with respect to numerous questions which were once the province of philosophers. A glance at the history of philosophy reveals a diminishing region of jurisdiction as the natural sciences have been successively hived off. The metaphor of philosophy as midwife to the sciences, abandoned after her services have been rendered, is an attractive one. Extrapolating from this grand narrative, resistance to the further scientisation of enquiry might seem destined towards deathly philosophy-of-the-gaps. But, even if this line of thought does something to motivate the naturalistic research programme, it requires careful scrutiny. Philosophers have of course advanced a number of arguments purporting to show why this or that field of research – modality, intentionality, mathematics, phenomenal consciousness, aesthetics – necessarily resists scientific treatment or reduction. We also find more ambitious attempts to show that science as a whole depends on particular metaphysical or epistemological assumptions. The scientist, it has been argued, has no choice but to take philosophical stands, either implicitly or explicitly, on matters such as fundamental ontology, modality or the nature of evidence.<sup>26</sup> I have no general quarrel with these types of arguments. The immediate problem, however, is that when it comes to a confrontation between naturalism and non-naturalism there are no commonly agreed rules of engagement. If we are to take seriously naturalistic claims concerning the omniscience of science, we should ask

---

reflecting relations in higher linguistic and cognitive processes, behaviourism lost credibility.’ (1990: 246). Because I identify philosophical naturalism with its methodological constraints, there is no room in my account for distinguishing between the philosophical theory and the metatheory.

<sup>25</sup> It may also be that the same items of knowledge are accessible via multiple avenues of enquiry. In this case it may make no difference which method is employed. Alternatively, other epistemic desiderata may be employed, such as simplicity, fecundity or conservativeness. Non-epistemic desiderata, such as beauty or moral worth, may also be appropriate considerations.

<sup>26</sup> For examples of this approach see Lowe (1998), Rea (2002) and Plantinga (2002).

how successful science is or could be at accounting for issues at the current core of philosophy. In line with the concerns of the present thesis, we need to consider whether naturalism has the resources to locate value-makers. In order to identify those resources, however, we firstly need to address the question of how to distinguish science from non-science in the sense relevant to the characterisation of EN. That will be the subject of the following chapter.

### **Chapter Three** **Naturalism and Scientific Practice**

In the previous chapter I argued that naturalism is best understood as an epistemological position which treats science as exclusively competent with respect to knowledge acquisition. The question then arises as to what it is that naturalists mean by ‘science’ as opposed to ‘non-science’, such that the former should be uniquely potent. Although this question is closely related to the classical ‘problem of demarcation’, it is not quite the same. I am concerned with finding a characterisation of science appropriate to naturalism and as it bears on a naturalistic approach to normativity. Just as in the previous chapter I considered characterising naturalism either in terms of its epistemological or ontological commitments, in section 3.1 I consider whether science should be characterised either in terms of its distinctive methods or its subject properties. I argue against both proposals and present a case in section 3.2 for taking refereed scientific articles as proxy indicators of scientific practice. Lastly, in section 3.3, I connect scientific papers to an objectifying stance which typifies scientific practice.

#### *3.1 The Properties and Methods of Science*

I would like to begin by considering the proposal that one can demarcate scientific disciplines from non-scientific disciplines in terms of their respective properties. The natural thought is that physics studies the physical, chemistry the chemical, biology the biological and so forth. More generally, whereas the natural sciences investigate natural properties, the subject properties of non-scientific disciplines are non-natural. This accords well with one of the intuitions which drives naturalism, namely that everything should be explicable in terms of the same sort of properties we require to explain paradigmatically physical objects, such as Austin’s medium-sized dry goods. However, this view of science would fit poorly with the arguments of the previous chapter, suggesting that naturalism could after all be given an ontological basis. Moreover, there are several independent reasons to think that it will not do to demarcate science from non-science in this way. One now familiar reason is the difficulty of saying what it is that distinguishes natural from non-natural properties. It would be circular to suggest that the former are those properties which figure in the theories of natural scientists. More significantly, we can point towards instances of two

subjects, one scientific and one non-scientific, with the same subject matter. Richard von Mises provides an interesting example:

The transition from green to brown *is* the river landscape, declares the landscapist Jacob van Ruisdael, and he accomplishes with this something similar to what the physicist does when he says that sounds are air vibrations. Every painting, every artistic creation is a *theory of a specific section of reality*. (1951: 303)

This is a curious view of art, and a rather implausible one when one considers non-representational artists whose work is intended to provoke a purely aesthetic response. Nevertheless, there is some plausibility to the claim that art can reveal something of the physical nature of the world, assuming we can make sense of non-propositional knowledge. If so, von Mises is right that art cannot be distinguished from physics on the basis that only the latter is concerned with representing physical properties. The important point, however, is that the overlap between scientific and non-scientific subject matter should be accepted by those engaged in naturalisation programmes. It was because von Mises was engaged in the project of unifying human knowledge under the umbrella of the physical sciences that it was important to him to show that art and physics have shared subject matter. Consider, for example, the project of naturalising epistemology. Even if one is persuaded by Quine that the traditional project of epistemology is bankrupt, why replace it with psychology? Why not pottery, dancing or physics? The transition from a discipline to its successor is a *non sequitur* unless the successor addresses a recognisably similar set of questions. It is for this reason that advocates of naturalised epistemology have been at pains to argue that psychology can answer the traditional epistemological questions concerning the conditions under which we are in a state of knowledge.<sup>1</sup> Thus, Quine maintains:

---

<sup>1</sup> The concern remains that rather than replacing traditional epistemology, naturalised epistemology merely institutes a shift from the normative question ‘what, if anything, can we have reason to believe?’ to the purely descriptive question ‘how do we arrive at our beliefs?’ I discuss this concern in chapter five and in Turp (2008).

A far cry, this, from the old epistemology. Yet it is no gratuitous change of subject matter, but an enlightened persistence rather in the original epistemological problem. (1974: 3)<sup>2</sup>

All of this suggests, then, that naturalists cannot distinguish science from non-science on the basis of their respective subject properties. This is just what we would expect given our earlier argument that naturalism is not an ontological thesis.

If science cannot be distinguished from non-science in terms of its subject properties, perhaps we should look to methodological differences. There may also be differences in the resultant theories, but for EN to be justified there must be some reason why scientific theories should be preferred to non-scientific theories. Most obviously, this preference could be explained in terms of the superiority of scientific method. Of course, questions concerning scientific method constitute a major topic within the philosophy of science, beyond the scope of the present thesis. Fortunately, however, we do not need to undertake a detailed analysis of the various rival theories. For, whatever the respective merits of, say, the hypothetico-deductive, Bayesian-inductive or inference-to-the-best-explanation models, EN does not depend on a formal account.

One reason why standard philosophical method-based solutions to the problem of demarcation are not germane to the purpose of characterising EN is that naturalists cannot be committed only to those methods which are *distinctive* of science as opposed to non-science. For EN to be defensible, scientific method must be understood widely enough to account for all obvious sources of knowledge. Much of our knowledge is available to apparently non-inferential, pre-scientific or extra-scientific consciousness. Examples include introspective statements such as ‘I know I exist’ or ‘I know I am in pain’, direct observational statements such as ‘I know Jack is sitting in front of me’, memory statements such as ‘I know I ate toast for breakfast’, or linguistic statements such as ‘I know what this sentence means’ or ‘I know that Jill said she would meet me for lunch’. One could easily multiply examples of such knowledge claims, which bear no

---

<sup>2</sup> Similarly, Goldman (1993) insists that scientific epistemology, if it is still to count as epistemology, must begin with an account of our ‘epistemic folkways’, by which he means our standard repertoire of epistemic concepts; justification, knowledge, rationality and the like.

obvious dependence on scientific method formally construed in terms of inferences from model experimental situations.<sup>3</sup>

None of this is to deny that scientists rely on these methods of knowledge formation in practice. Of course, successful science presupposes our ability to reliably acquire knowledge in ways not unique to science. As scientific practice extends to methods employed in common with other disciplines, and with no discipline in particular, so, then, should the appropriate characterisation of scientific methodology. An account solely in terms of what is distinctive of scientific method will not do. This must be the case for the epistemological naturalist to maintain that science is the one true path to knowledge. However, we then return to the question of why those shared methods should be suitable for knowledge acquisition in a way that methods employed solely in non-scientific contexts are not. Given the apparent heterogeneity of distinctively scientific methods and methods shared with other disciplines, it might seem unlikely that they share a common property other than that of being employed by scientists. However, EN requires that only beliefs arrived at through scientific methods are proper candidates for knowledge. Thus, scientific methods must share some normative property in virtue of which they are uniquely suited to generate justified beliefs. The bare fact that such methods are employed by scientists, unmotivated by other considerations, is an implausible candidate.

For EN to be true, then, there must be something special about the methods employed by scientific practitioners, something to be captured in normative terms. Moreover, it seems that the appropriate normative account must be drawn from scientific practice, as opposed to the philosophy of science, insofar as the latter is a distinct discipline employing distinct methods. This provides us with another reason to doubt that standard methodological solutions to the problem of demarcation will help to characterise EN. For, it would be dialectically counterproductive to insist on characterising the naturalist's position employing

---

<sup>3</sup> The psychologist Kevin Dunbar goes so far as to argue that there are *no* cognitive strategies unique to science, the difference between science and non-science resulting from the sequence in which various cognitive strategies are deployed: 'Without question, human beings, including children, possess each of these cognitive skills. The way that these skills are brought together is what makes the thinking of a baseball player different from a molecular biologist.' (2002: 168)

philosophical or normative analyses which transcend actual scientific practice. The naturalist will not accept an external criterion based on methods she does not accept. It is not the business of philosophy, she will insist, to tell scientists what they ought to be doing, or whether what they are doing counts as scientific. Indeed, it is something like this thought which motivates EN in the first place. Quine writes that naturalists have ‘the robust state of mind of the natural scientist who has never felt any qualms beyond the negotiable uncertainties internal to science’ (1981: 72).<sup>4</sup> If ‘robust’ has a non-rhetorical function, it is likely that it refers to the naturalistic refusal to engage with traditional philosophy discussed in the previous chapter. If we want to know what the naturalist considers special about science, we need to look to scientific practice, not to prior philosophy.

### *3.2 Articles as Proxy Indicators of Scientific Practice*

We come, then, to the idea that the most promising way to characterise naturalism is as a commitment to the omnicompetence of a research programme governed by the norms internal to scientific practice. How though are we to identify these norms? Unfortunately, they cannot be straightforwardly read off from scientific practice. Whilst scientists sometimes contaminate samples, distort data and forget crucial facts, such activities do not count as norm-governed scientific practice. This is not only because such behaviours are statistically uncommon. Were a majority of working scientists to conduct their research by smashing culture tubes and divining theories from the scattered shards, culture tube smashing would not thereby count as scientific. Culture tube smashing violates the internal norms which are partly constitutive of scientific practice. One is not a scientist at all if one deviates too often from these norms, inculcated over a lengthy apprenticeship.

Legislating and policing these internal norms is a matter for scientific practitioners, and describing them is most properly a matter for science studies and ‘lab anthropology’. Providing a detailed descriptive account of scientific

---

<sup>4</sup> Similarly, Penelope Maddy describes ‘the fundamental spirit that underlies all naturalism’ as ‘the conviction that a successful enterprise, be it science or mathematics, should be understood and evaluated on its own terms, that such an enterprise should not be subject to criticism from, and does not stand in need of support from, some external, supposedly higher point of view’ (1997:184). Again, Ritchie advocates that we ‘start thinking about science within scientific practice, not trying to provide a justification of science from outwith science’ (2008: 88).

practice is not the task of philosophers *per se*. Moreover, the fact that we do not find any such account in the writings of self-proclaimed philosophical naturalists seems to indicate that it is not essential to the characterisation of their position. On the other hand, if we are to advance our understanding of the relationship between naturalism and the metaphysics of normativity, it is essential to do something more substantial than gesture towards scientific practice. In lieu of a detailed descriptive account, I propose, therefore, that we take the refereed scientific journal article as a proxy indicator of scientific practice. I shall spend the remainder of this section defending the proposal, before drawing out certain relevant consequences in section 3.3.

Refereed articles are the premier form of knowledge dissemination within the scientific community and, as such, subject to the strictest standards of peer review. Indeed, it is the purpose of the peer review process to ensure that the research under review conforms to the professional norms governing the discipline. As Fred Suppe writes:

The principal tasks of a scientific paper are to present knowledge claims and support them with sufficient explicit justification to enable discipline members and gatekeepers to evaluate whether to accept these claims and admit them into the discipline's domain of putative knowledge. (1998: 384)

Scientific articles are public and open descriptions of research, which have been judged by authoritative scientific practitioners to conform to the norms of their discipline. I am not proposing that a biconditional holds between successful scientific practice and published research in refereed journals. Obviously enough, publication is not sufficient.<sup>5</sup> Reviewers are fallible and consequently a certain amount of pseudoscience slips past the gatekeepers into print. This is perhaps most likely in journals with a relatively poor reputation, and also in certain prestigious letters journals in which truncated pieces can be rushed to press ahead of competitors. As this is predictable to a certain degree, we might

---

<sup>5</sup> Nor is it necessary. It would seem arbitrary to bar research as scientific on the grounds that a researcher leaves it absent-mindedly in a drawer as opposed to an editor's inbox. Perhaps the more pertinent point is whether the research would be accepted for publication, were it to be submitted. However, this type of modal claim seems far more problematic for a naturalist wanting a criterion involving only scientifically discoverable properties.



make the proposal more precise by counting only papers published in journals with a high impact factor, or papers of a certain length, or, retrospectively, papers with a certain number of positive citations. However this may be, the general principle is that non-science or pseudoscience ought not to appear in refereed journals. If non-science appears, it is by error. Moreover, the error of allowing the non-science into print is to be attributed to the reviewers and is to be corrected by their peers. No appeal is made to an external philosophical tribunal. If the alleged research findings are important within the field, it is likely that they will be rebutted in future papers, providing a corrective mechanism internal to the practice.

The primary advantage of taking refereed articles as proxy indicators of scientific practice is that it allows us to provide an account of the normative standards governing scientific practice without relying upon any prior philosophy. Although a normative realist might object that it conflates the norms which scientific practitioners take to govern their discipline with the norms that in fact govern it, naturalists would reject the presupposition of a two-level framework. On a naturalistic account, there is no normative standard external to scientific practice to be conflated with the norms internal to science. Similarly, naturalists will not be in a position to object that the proposal is circular, for any account of science without recourse to extra-scientific resources must be circular. Therefore, if a naturalist were to disagree with my constructive proposal, it could only be on the grounds that scientific papers misrepresent the norms internal to scientific practice. Before addressing this type of objection, we might note two further advantages of taking refereed scientific articles as our basis for identifying scientific norms. First, the proposal chimes well with our earlier conclusion that naturalism concerns methodology not ontology. A scientific paper is selected for publication depending upon the cogency of its reasoning, rather than its results. When science progresses, the results are not known in advance of the research.<sup>6</sup> A second advantage of leaving the task of judging

---

<sup>6</sup> This is not to suggest that scientists have no advance beliefs concerning the outcome of their experiments. Indeed, experimental results are sometimes considered so unlikely that they are rejected or ignored without any methodological flaw being identified. As Max Perutz observes, ‘an experiment is an experiment and calls for an explanation, but it can be hard to discover the explanation for someone else’s spurious results, and without it they cannot be convincingly

what counts as scientific to the relevant gatekeepers is that it respects the clearly permissible differences in methodology between disciplines. Particle physicists judge what counts as legitimate method within particle physics, molecular biologists within molecular biology, and so forth. Thus, we can circumvent the potential objection to EN that the sciences are too diverse to exhibit a single methodological structure.<sup>7</sup>

It might still be objected that taking refereed articles as proxy indicators of scientific practice is likely to mislead us insofar as there is a mismatch between what scientists do and what they say that they have done. For instance, Schickore writes:

Since the research reports [within published scientific articles] do not preserve the temporal order or logic of the thinking processes nor the steps of the experimental pathway, they do not provide any dependable information about the nature of scientific practice. (2008: 330)<sup>8</sup>

Whether or not one agrees that research reports provide *no* dependable information, it is true that scientific articles are not straightforward reports of research undertaken. This should come as no surprise, for, as already mentioned, the purpose of a scientific paper is to provide evidence-based justification for knowledge claims. We should be no more taken aback by the fact that scientific papers are not transparent reports of scientific practice than by the fact that a philosophy paper is a poor record of its author's thought processes. In both cases we are not being offered a record at all, but the marshalling of reasons in favour of knowledge claims. Much of a scientist's time is taken up pursuing lines of research which turn out to be irrelevant to the justification of her final claims, or which she misconstrues the significance of at the time. In the context of justification the temporal order in which the evidence was obtained, or the temporal order in which the thoughts occurred to the author, is simply irrelevant. One way to make this clear is by noting that the term 'science' is ambiguous between activity and product. So, we can just as properly say that an experiment

---

disproved' (1998: 138). It remains the case, however, that experimental data properly arrived at ought to trump prior expectations.

<sup>7</sup> See, for instance, Chalmers (1999: Ch. 11).

<sup>8</sup> Similarly, Peter Medawar writes that 'it is no use looking to scientific "papers", for they not merely conceal but actively misrepresent the reasoning that goes into the work they describe' (1961: 169).

is science as we can say that the resultant theory is science. Epistemological naturalists, however, are committed to a view concerning knowledge, which is a scientific product. Although naturalists will point out that the product is justified by the activity, it is a truism that scientific activity can only justify a knowledge claim when that activity is relevant to its justification. Insofar as scientific articles represent successful attempts to filter out of reports research activity which is irrelevant to the justification of the knowledge claims, they thereby filter out information which is also irrelevant to the naturalists' claims. In this way, scientific articles are better indicators of the scientific norms relevant to the justification of EN than a perfectly transparent research report would be.

It must be conceded, however, that even if an article serves as a good indicator of norms, it is not an explicit statement of them. It is no more the role of scientific articles to articulate the norms by which they are governed than it is the role of a game of chess to articulate its own rulebook. Moreover, the fact that certain norms or methodological standards are reflected in a scientific article does not entail that the research was conducted in consultation with those norms. Would it not be better, therefore, to look directly at the scientific rulebook, to explicit statements of scientific method? One difficulty is that scientific method is not itself part of the subject matter of the natural sciences. It is the task of biologists to study living matter, not to study the general features of the methods that they employ to study living matter. When scientists do write broadly concerning methodology it is generally for the benefit of their students. Most common is a brief three or four page discussion at the beginning of a textbook. Unfortunately these discussions rarely describe science as it is actually practised, most frequently garbling ideas from the philosophy of science instead.<sup>9</sup> Indeed, the overwhelming impression is that most scientists have little serious interest in providing rigorous formal accounts of the methods they employ, or doubt the possibility of doing so. The following comments come from philosophically reflective scientists, both Nobel laureates:

There is indeed no such thing as “the” scientific method. A scientist uses a

---

<sup>9</sup> Blachowicz (2009) analyses 70 textbook discussions of scientific method, and has a certain amount of fun highlighting their inconsistencies and errors.

very great variety of explanatory stratagems, and although a scientist has a certain address to his problems – a certain way of going about things that is more likely to bring success than the gropings of an amateur – he uses no procedure of discovery that can be logically scripted. (Medawar 1984: 51)

Scientists rarely follow any of the scientific methods that philosophers have prescribed for them. They use their common sense. (Perutz 1998: 123)

Of course, the selection of appropriate scientific strategies is not really a matter of ‘common sense’, but rather a matter of the sense common to scientific practitioners who have undergone a standardised process of training. Thus, Polanyi emphasises scientists’ reliance on ‘tacit knowledge’, the kind of unarticulated wisdom which allows an experienced physicist to see subnuclear events represented in a bubble chamber photograph where a novice would see only confusion.<sup>10</sup> The importance of tacit knowledge, he maintains, means that ‘the methods of scientific inquiry cannot be explicitly formulated and hence can be transmitted only in the same way as an art, by the affiliation of apprentices to a master’ (1962: 69). A scientist’s possession of procedural knowledge does not entail possession of the corresponding propositional knowledge.

The above is of course no demonstration of the impossibility of articulating a formal account of scientific method, although it does perhaps cast doubt on the project. Rather it is to emphasise that scientific practitioners do not consciously rely on any such account and cannot be depended upon to provide one. Moreover, the norms governing scientific practice exhibit considerably greater plasticity than we see in a strictly rule-governed practice such as chess. We need not say with Feyerabend that ‘anything goes’ in order to acknowledge that norms governing scientific method are related to the pursuit of changing scientific goals.<sup>11</sup> Scientists do not respect these norms for their own sake, but because

---

<sup>10</sup> See further Kuhn (1970: 111).

<sup>11</sup> Just what the aims of science are is of course a matter of considerable controversy. Popular candidates include truth, verisimilitude, knowledge, explanation, prediction, puzzle-solving and empirical adequacy. Resnik (1993), however, argues that science is without aims. The core of his argument is that scientific aims must either be a matter of the goals scientists in fact have, or a normative matter to be settled by philosophers of science. But, as a descriptive matter, there are no aims which are shared by all and only scientists (and nor does science have an appropriate hierarchy to attribute aims to a type of corporate structure). Moreover, Resnik argues that the normative pronouncements of philosophers float free of scientific practice and are therefore too abstract and general to be of interest to working scientists. Whatever the merits of his particular arguments, Resnik is correct in his insistence that an account of scientific practice should meet the dual desiderata of reflecting what it is that scientists actually do and holding their activities to

experience indicates that they lead to the satisfaction or partial satisfaction of scientific aims. Thus, scientific methodology allows for the innovation and change that modifies but does not destroy the practice.<sup>12</sup> We might think, for instance, of the move away from the direct evidence of the senses to the evidence of instruments such as telescopes, the debate between Galileo and his Aristotelian peers being a much discussed example.<sup>13</sup> Although scientific journals are essentially conservative, it is reasonable to suppose that they respond to changes in scientific practice more quickly and accurately than textbooks.

### *3.3 Scientific Articles and Objectification*

I have proposed, then, that we take refereed scientific articles as proxy indicators of norm-governed scientific practice. In the course of the argument, I addressed two potential objections, namely (i) that scientific papers misrepresent scientific practice and (ii) that scientific papers are worse indicators than scientist's explicit methodological statements. Earlier I rejected the view that scientific practice can be characterised in terms of its ontological commitments or in terms of an atemporal methodology. As the norms governing scientific papers are related to the pursuit of changing proximate goals, acceptable methods vary across scientific disciplines and over time. Nevertheless, I think it is possible to discern certain essential features of science reflected in the norms governing scientific publication. Of particular significance, as I shall now explain, is the objectifying character of the scientific stance.

In a guide to scientific publication, primarily written for the benefit of junior research scientists, Robert Day states that a paper must fulfil the following criteria:

---

normative standards. The present move towards scientific practice, with refereed articles as success indicators, provides, it seems to me, the appropriate combination of normative and descriptive criteria.

<sup>12</sup> As Polanyi writes: 'The professional standards of science must impose a framework of discipline and at the same time encourage rebellion against it. They must demand that, in order to be taken seriously, an investigation should largely conform to the currently predominant beliefs about the nature of things, while allowing that in order to be original it may to some extent go against these. Thus, the authority of scientific opinion enforces the teachings of science in general, for the very purpose of fostering their subversion in particular points.' (1962: 58–9)

<sup>13</sup> See, for example, Feyerabend (1975) and Chalmers (1999: 163–8).

Sufficient information must be presented so that potential users of the data can (i) assess observations, (ii) repeat experiments, and (iii) evaluate intellectual processes. (1979: 2)<sup>14</sup>

On a reasonably broad interpretation, (i) and (iii) are simply requirements of clarity and transparency in content and logic respectively. We should expect this from philosophy as much as from science. Let us, then, concentrate on (ii), the requirement of reproducibility. As Day emphasises later:

The cornerstone of scientific method *requires* that your results, to be of scientific merit, must be reproducible ... The fact that experiments are unlikely to be reproduced is beside the point; the potential for producing the same or similar results *must* exist. (1979: 26)

The deontological terms '*requires*' and '*must*' – Day's italics – can be taken to indicate that we are dealing with necessary conditions for engaging in scientific practice, rather than recommendations related to local scientific goals.<sup>15</sup> If one's research is not in principle reproducible, it is not scientific. In order to meet the demand of reproducibility, however, agents within a scientific research programme must adopt a particular epistemic stance towards the world, excluding certain putative methods of knowledge acquisition and sources of content. One distinguishing feature of the scientific stance is that it is objectifying. The sense of 'objectifying' I have in mind refers to the distinction between objects and persons, although the distinction between objectivity and subjectivity is a close conceptual relation. In adopting an objectifying stance, a scientist describes and acts upon the world as if it were wholly populated by disenchanted objects, standing in value-free, causal and constitutive relations,

---

<sup>14</sup> Day adopts these criteria from a definition of a scientific paper proposed by the Council of Biology Editors, now the Council of Science Editors, a professional organisation for editors of scientific journals.

<sup>15</sup> There is of course a sense in which philosophy is also reproducible. At least, philosophers are generally able to follow one another's lines of thought. On one view, this is a matter of entertaining a series of numerically identical propositions. The trouble is that even when philosophers follow the same line of thought, they frequently do not agree on its significance or non-formal implications. Equally, philosophers often agree concerning the validity of an argument, but not its soundness. This is rather as if a scientists successfully reproduced the methods sections of each other's papers, but with no consistency in terms of results. One possible explanation is that philosophical questions do not, by and large, allow for the control of variables. Related to this, there is no clear limit to what is potentially relevant to a philosophical argument, or at least no generally agreed-upon procedure for determining what is relevant. As a result, individual philosophers appeal to those arguments and features of the world which strike them as most salient.

potentially amenable to experimental intervention.<sup>16</sup> Employing Sellars' distinction, it is a turn from the manifest image to the scientific image of man-in-the-world, with things not persons as its objects. As persons are objectified, they are no longer recognised as *sui generis* elements of the world, but reduced to biological or perhaps psychological processes. It is this approach which, seeking to demystify the mind, generates familiar philosophical issues such as the problem of intentionality and the hard problem of consciousness.

The objectification of the perceptible world is a familiar element of the Modern Scientific Worldview and leads more or less immediately back to the problem of locating value-makers discussed in chapter one. It is difficult to see how the constituents of a disenchanted world could have the requisite normative force to guide thought or action. Normative properties begin to look queer. Instead of labouring this point further, however, I would like to turn to look at another aspect of the scientist's objectifying stance. This is the idea that for scientific work to be reproducible, the elimination of persons in favour of objects must be carried through consistently. In particular, the scientist is himself a person with a unique point of view on the world, and so must make every effort to remove himself from the scene of his investigations. Looking at scientific writing, this aim is reflected in the convention that papers are written in the third person, despite the research having been conducted by particular scientists in particular settings.<sup>17</sup> It is also reflected in the norm of avoiding indexical expressions. As Russell writes:

No egocentric particulars occur in the language of physics. Physics views space-time impartially, as God might be supposed to view it; there is not, as in perception, a region which is specially warm and intimate and bright, surrounded in all directions by gradually growing darkness. A physicist will not say "I saw a table", but like Neurath or Julius Caesar, "Otto saw a table"; he will not say "A meteor is visible now", but "A meteor was visible at 8h. 43m. G.M.T.", and in this statement "was" is intended to be without tense. (1962: 102)

---

<sup>16</sup> Following a decade of observing laboratory meetings, Dunbar reports that 'causal reasoning is one of the key types of cognitive activity that ... scientists engage in – accounting for over 80% of the statements made at a meeting' (2002: 157).

<sup>17</sup> The convention is of course disingenuous and seems to be falling out of fashion. *Nature* has for some years now cultivated the editorial policy of encouraging the use of the active voice and consequent first person (usually plural). *Science* has more recently followed suit.

As there is only one here and now, at the time and place of utterance, indexical elements cannot be reproduced.<sup>18</sup> *Pace* Russell, the ‘was’ of ‘a meteor was visible at 8h. 43m. G.M.T.’ should be tensed, but only as a temporary matter of publishing etiquette. At the time of writing, a scientific paper has not been published and so its contents have not yet been accredited as scientific knowledge. When scientists refer to previously published work, including their own, they employ the present tense in order to indicate that they are referring to timeless knowledge. For instance, it would be correct to write “x upregulates y (Smith 2010)”, but incorrect to write “Smith (2010) found that x upregulated y”.

Although superficial features in themselves, these stylistic conventions are indicative of the self-effacing nature of the scientific stance. The cumulative result is the impression of a paper’s author fading to an extensionless point.<sup>19</sup> Thus, Popper asserts that “‘scientific knowledge’ may be regarded as subjectless’ (1970: 57). Although one might doubt whether there is much literal sense to be made of subjectless knowledge, this does appear to be the regulative ideal governing scientific practice and made manifest in scientific papers. At least, the presentation and interpretation of research in a scientific paper should not depend upon the idiosyncrasies of a particular subject. Consider, for instance, Day’s comment:

The preparer of a scientific paper is not really an “author” in the literary sense. In fact, I go so far as to say that, if the ingredients are properly organized, the paper will virtually write itself. (1979: 4)

It is revealing in this respect that science departments sometimes hire third-party specialists for the purpose of writing up experiments, a situation predicated upon the interchangeability of one ‘preparer’ with another. Whatever the egocentric particulars of the researcher, and however vital these may have been in the

---

<sup>18</sup> Some philosophers have questioned the intelligibility of adopting an indexical-free stance. The concern is that agency requires a view from somewhere. Thus, McGinn writes that ‘when I imagine myself divested of indexical thoughts, employing only centreless mental representations, I eo ipso imagine myself deprived of the power to act’ (1983: 104). For the original statement of the problem of the essential indexical see Perry (1979).

<sup>19</sup> Whilst in the biological sciences there remain conventions enabling one to discern something of the respective inputs of the named authors on a paper, high-energy experimental physics has reached the point at which authors are listed alphabetically. This can run to several hundred names at large collaborations such as CERN. For detailed discussion see Knorr Cetina (1999: 167).



context of discovery, they cannot be replicated at will and are therefore excised from the published report.

Moreover, just as the organisation of the ingredients should not depend on the peculiarities of persons, neither should the production of the ingredients. Whilst skilled, the physical work of experimentation in modern laboratories is highly automated and frequently devolved away from group heads to doctoral or postdoctoral students. For instance, a biologist's contact with her subject matter is normally mediated by gel apparatuses, PCR (polymerase chain reaction) machines, electron microscopes and the like. The resulting data is then processed using standard analytical tools, such as Southern blots and western blots, once more automated. In large pharmaceutical laboratories cell cultures are mechanically exposed to millions of biochemical compounds whilst a computer programme 'searches' for significant patterns. The researcher only becomes involved after the computer 'discovers' something of potential interest. Although it is true that the automation of research is motivated by considerations of speed and efficiency, it also reflects once again the regulative ideal of subjectless knowledge. As Joseph Hanna writes, 'science makes progress through the replacement of subjective protocol judgments by objective artifactual data' (2004: 342; emphases removed). As with the objectifying stance quite generally, automation has the significant epistemic benefit of promoting distributed cognition.<sup>20</sup> Individual scientists become more or less interchangeable nodes in a cognitive network, collectively performing more work as a result. It helps to foster what Dewey described as the 'cooperative tendency towards consensus which marks inquiry in the natural sciences' (1929: 30).

Given the account of normative epistemology I shall be developing in chapter nine, it is also worth highlighting that whilst emotional responses play a significant role in the context of discovery, they are excluded from scientific

---

<sup>20</sup> See Giere (2006) for a detailed discussion of the Hubble Space Telescope programme as an example of distributed cognition.

articles.<sup>21</sup> As well as the distorting effect emotions can have upon cognition, emotional responses are unlikely to be consistently repeated and confirmed, and unlikely, therefore, to receive experimental backing. One reason for the disturbing influence of the emotions is their function of highlighting aspects of the world as salient and valenced. As the perceived salience of some feature depends on contextual factors and background interests, the objectifying stance cannot take note of emotional responses. Borrowing a phrase from the particularism debate, we are left surveying a flattened normative landscape.

I have argued, then, that the rhetoric of scientific papers suggests a regulative ideal of subjectless knowledge, which is one facet of an objectifying stance that takes things rather than persons as its objects. This stance enables reproducibility; an essential feature of scientific practice which facilitates distributed cognition. It might be objected, however, that this is a misleading caricature of scientific practice. For instance, methods sections rarely provide sufficient detail to replicate an experiment precisely and it is in any case uncommon for an experiment to be replicated. Pragmatically, there is little point. If the results match, the research will not normally be interesting enough to publish. If the results do not match, this can be attributed to minor methodological differences. Protocols invariably underdetermine action, and no two laboratories are identical. Perhaps more seriously, it might be objected that scientists are not interchangeable *automata* and that it is fortunate that they are not, given that science is sometimes driven forwards by radical shifts of perspective initiated by individuals. This is a phenomenon we associate with great scientists such as Copernicus, Darwin and Einstein, men who lived in different worlds according to Kuhn. Yet, even if a psychological or historical explanation of their advances would make reference to the genius peculiar to these men, their theories were not *justified* by their peculiarities. Indeed, I have already agreed that the norms governing scientific publication do not reflect scientific practice. This is because the norms relate to the context of justification, not to the context of discovery. The context of justification is the appropriate place to look for naturalistically acceptable norms given that

---

<sup>21</sup> For a discussion of the role of affective states in scientific research see Hookway (2002) and Thagard (2002).

naturalism amounts to a claim concerning scientific knowledge. If subjectless knowledge is a regulative ideal governing the context of justification, it is no objection that this is not fully reflected in the context of discovery.

Even if it is granted that science is appropriately characterised in terms of its objectifying stance, one might nevertheless protest that this is an incomplete characterisation. And indeed, this would be true.<sup>22</sup> For I have not attempted to provide a solution to the problem of demarcation in the form of necessary and sufficient conditions. In fact, I would treat such a project with suspicion given the non-rigid identity conditions of scientific practice. Rather, I have attempted to identify the most promising approach to constructing a naturalistically acceptable account of science and to provide that account with substantive content relevant to the programme of developing a naturalistic account of normativity. In the previous chapter I argued that naturalism is a commitment to scientism. I have now proposed that a naturalistically acceptable account of science could profitably take refereed scientific articles as proxy indicators of the normative standards governing scientific practice. Thus, naturalism becomes the view that all knowledge is of the sort which could in principle meet the peer review standards of scientific journals. All knowledge must be acquirable from the perspective of an objectifying stance which effaces the subject and disenchant the world. I shall begin to consider the prospects of such an account in the next chapter by looking at the concrete proposal of naturalising epistemology and the manner in which proponents of naturalised epistemology propose to account for epistemic normativity.

---

<sup>22</sup> One obvious omission from my discussion is the putatively empirical character of science. In fact, scientists are more likely to be concerned with the question of whether research is experimental than with whether it is empirical. However, these questions appear to collapse when we notice that empirical knowledge is most plausibly understood in terms of the existence of a causal chain linking an object and a mental state. As experimentation requires interaction, possibly indirect, with some part of that causal chain, it is empirical. However, it might be better not to insist on characterising science in empirical terms given naturalists' reliance on the formal disciplines of logic and mathematics.

## **Chapter Four** **Naturalised Logic and Normativity**

In chapter two I argued that naturalism is best construed in terms of its epistemological commitments. In chapter three I offered an account of the distinction between science and non-science. Whether or not my constructive proposal for a naturalistic solution to the demarcation problem is the best available, my claim that an objectifying stance is a regulative ideal of scientific practice should appear independently plausible. It is in part a function of the public nature of science and its methodological dependence on distributed cognition. I also argued in chapter two that the most productive manner in which to judge naturalism is according to its fruits. In the next two chapters I shall take up this task and examine the resources naturalists might bring to bear on the problem of explaining normativity. In chapter five, I shall consider the popular strategy of reducing normative reasons to instrumental reasons. First, however, I propose to look closely at Penelope Maddy's project of constructing naturalistic foundations for logic. I shall begin in section 4.1 by sketching out her account. This will provide us with a concrete example of naturalised philosophy. In section 4.2 I shall consider whether Maddy's naturalised logic contains within it an explanation for the normative force of logical inference. After distinguishing between explanatory and justificatory reasons, I conclude that it does not. In section 4.3, I take a look at logical psychologism and consider whether it is possible to bridge the gap between explanations and justifications. I conclude that an independent normative standard is required and lead into the next chapter by considering Maddy's view in relation to instrumentalism.

### *4.1 Maddy's naturalised philosophy of logic*

I argued in chapter two that it is dialectically counterproductive to insist on assessing the cogency of naturalism in an extra-scientific, philosophical tribunal. Naturalists reject any such two-level system, and should not attempt to mount a philosophical defence of their position.<sup>1</sup> For, as we have said, if naturalism is a

---

<sup>1</sup> This dialectic has not of course prevented self-proclaimed naturalists illegitimately advancing a number of philosophical arguments. See Almeder (1998: Ch. 1) for synopsis and criticism of the most common of these.

thesis, it is a philosophical thesis and thus only directly justifiable by methods which the naturalist rejects. Nevertheless, I have urged a charitable interpretation, highlighting parallels with the epistemic position of theorists within the philosophical research tradition as viewed from a theistic perspective. Rather than defending naturalism directly, naturalists should concentrate on pursuing their research programme and offering illuminating accounts of traditional philosophical problems. Perhaps in that way they will have no need of philosophical hypotheses.

In this light, Penelope Maddy (2007) offers the most fully developed and consistent naturalistic attempt to resolve traditional philosophical questions that I know of. One aspect of her discussion is the invention of a character who has fully taken to heart Quine's rejection of first philosophy. This character, the 'Second Philosopher', is 'born native to the laboratory' (2007: 91) and draws all her inspiration from her birthplace.<sup>2</sup> Consequently, she 'doesn't speak the language of contemporary science "like a native", she *is* a native' (2007: 308). Her warrant for belief in some proposition just is the scientific evidence in favour of that proposition.<sup>3</sup> Thus, Maddy draws our attention to 'the distinction between "I believe in atoms because I believe in science and it supports their existence" ... and "I believe in atoms because Einstein argued so-and-so, and Perrin did experiments such-and-such, with these results"' (2007: 85-6).<sup>4</sup> When the Second Philosopher justifies her beliefs, it is on the basis of scientific data, not on the basis of more fundamental epistemological principles. The only epistemic norms she accepts are those internal to scientific practice. She is lost if

---

<sup>2</sup> As it is sometimes archly remarked, Quine's writings are more suggestive of inspiration from the library than from the laboratory. See Fogelin (1997: 561).

<sup>3</sup> Thus, Maddy's position seems to have immunity against Plantinga's Evolutionary Argument Against Naturalism (1993, 2002), which I look at in section 5.4 below. According to Plantinga, naturalism and evolution form an unstable coalition. If evolutionary theory is true then our cognitive faculties have been selected for fitness, rather than designed for reliability with respect to truth. This constitutes a defeater for the belief that our cognitive faculties are reliable, and so a defeater for our belief that naturalism is true, given that this belief was formed by cognitive faculties which we cannot generally take to be reliable. In response, it seems that Maddy could deny that we need any general warrant for the belief that our cognitive faculties deployed in the scientific context are reliable.

<sup>4</sup> We find a similar point being made by Dewey: 'If the empirical method were universally or even generally adopted in philosophizing, there would be no need of referring to experience. The scientific inquirer talks and writes about particular observed events and qualities, about specific calculations. He makes no allusions to experience; one would probably have to search a long time through reports of special researches in order to find the word.' (1929: 2)

asked for external, non-scientific justification and she consequently declines to engage in second-order debates concerning the epistemic framework within which she operates. She attempts no general defence of scientific method other than evaluating its conduciveness to the goals internal to scientific practice. If she is guilty of dogmatism, it should also be noted that the Second Philosopher can do no other than appeal to her own methods, for she recognises the validity of no others. As Maddy writes, ‘the Second Philosopher has no perspective but her own to offer’ (2007: 312).<sup>5</sup> In brief, Maddy’s Second Philosopher adopts precisely the epistemic standpoint I have been urging upon clearheaded naturalists.

Let us, then, consider a concrete example of Maddy’s Second Philosopher at work, namely the project of constructing a naturalised account of elementary logic. Logic is normative if anything is, and so a fully naturalised account of elementary logic is *ipso facto* a naturalised account of one aspect of normativity. Although there is more technical detail in Maddy’s account than presented below, my primary concerns are with the general structure, method and outcome. Essentially, Maddy attempts to defend three claims, which she takes to provide the basis for a philosophy of logic. They are as follows:

- (1) Logic is true of the world because of its underlying structural features
- (2) Human beings believe logical truths because their most primitive cognitive mechanisms allow them to detect and represent the aforementioned features of the world
- (3) Human beings are so structured cognitively because they live in a world that is so structured physically. (2007: 226)

Thus, we have claims concerning (1) the truth conditions of logic and (2) the epistemology of logic. We also have in (3) a principle connecting (1) and (2).

Maddy’s first step in defence of this account is to describe a simplified, abstract ‘KF-world’ consisting of elements corresponding to the elements of a

---

<sup>5</sup> Although, it is worth adding, naturalists do have recourse to intra-scientific criticism. A critique of one scientific discipline can be mounted from within another. See further Roland (2007: 430-3) who discusses the possibility of naturalistic intra-scientific criticism under the rubric of ‘disciplinary holism’.

rudimentary system of logic.<sup>6</sup> In particular, A KF-world contains objects,  $a, b, c, \dots$ , which bear various properties,  $P, Q, \dots$ , and stand in various relations with various numbers of arguments,  $R, S, \dots$  (2007: 228). From an object failing to enjoy some property or stand in some relation, we have the idea of negation. From an object standing in relation to more than one other object, or being related either this way or that, we have the ideas of conjunction and disjunction respectively. Also, properties can hold universally in a KF-world, with existentials accompanying universals in the standard way (i.e.  $\forall xPx \leftrightarrow \neg\exists x\neg Px$ ). Maddy discusses the interconnections between these elements in considerable detail, explaining how situations in the KF-world correspond to logical connectives, and indicating how this world differs from an alternative world modelling intuitionistic logic.

With this model in hand, Maddy argues that the actual world contains elements corresponding to those of a KF-world. Beginning with objects, Maddy argues that ‘common sense clearly endorses the idea that the world contains many medium-sized physical objects’ (2007: 234). As well as common sense, scientific discourse makes apparently ineliminable reference to a great many objects, from cells to stars. Physics also provides an account of objects as complicated arrangements of atoms held together by atomic bonds, distinct from neighbouring atoms and regions of space, and generating electromagnetic fields. The Second Philosopher is not interested in, and indeed can make no sense of, the sort of hyperbolic doubt concerning the existence of objects which would undermine all of her methods. So, there is no room for scepticism concerning the existence of macroscopic objects.<sup>7</sup> Again, the fact that objects enjoy properties and stand in relations is both reflected in scientific discourse and a straightforward matter of common sense (2007: 237–9). If it seems laboured, the

---

<sup>6</sup> The designation ‘KF’ derives from the historical roots Maddy traces in the work of Kant and Frege, but nothing substantial turns on the terminology.

<sup>7</sup> At the subatomic level matters become more complicated, and Maddy sides with those who argue that objects, properties and relations break down at this point (2007: 239–40). She also argues that there is vagueness and indeterminacy through and through (2007: 240–4). As a consequence, Maddy modifies (1) to ‘(1’): rudimentary logic is true of the world insofar as it is a KF-world, and in many but not all respects, it is’ (2007: 244), and also modifies (2) and (3) accordingly. When K-F structuring is not present, Maddy argues, our logical intuitions are likely to fail us (2007: 272). Nevertheless, the direct objects of human cognition are at the macro level. Although these complications are important, they do not affect the thrust of her account in a way that is relevant to the concerns of this chapter.

purpose of this defence is to show that KF-structures are naturalistically respectable. There is no call for a transcendental or otherwise spooky analysis. There is also no recourse, explicitly at least, to prior metaphysical analyses of the relevant ontological categories of objects, properties and relations. Overall, the presence in the actual world of elements corresponding to the elements of a pared down logic supports the view that the former provides the truth conditions for the latter.

In defence of claim (2), concerning the logical structure of cognition, Maddy draws on findings from developmental psychology.<sup>8</sup> For instance, there is strong evidence that infants are able to individuate objects from a very young age. In particular, it seems that infants as young as three months are able to pick out so-called ‘Spelke objects’. These are connected and bounded regions of matter that move as one.<sup>9</sup> Factors such as similarity, continuity and regularity of form are less important than the ‘common fate’ of the parts of the object when it moves. Based on Spelke’s research, Maddy writes:

It seems humans are so configured, biologically, that they come to perceive a world of Spelke objects, without instruction, given ordinary maturation in a normal environment. Or, in the Second Philosopher’s terms, the ability to perceive Spelke objects is part of a human being’s most primitive cognitive equipment. (2007: 258)

Building on the capacity to recognise Spelke objects, a grasp of the criteria of individuation associated with sortal concepts is acquired from around twelve months. This is correlated with, and may depend upon, language comprehension of the relevant count noun.<sup>10</sup> It takes a certain level of conceptual sophistication

---

<sup>8</sup> Caution is in order when interpreting (2) because ‘primitive’ is ambiguous. Maddy’s intended thought is presumably that the primitive cognitive mechanisms are deep because our mature cognitive mechanisms are built upon them. This is in a manner similar to, say, emotional consciousness depending upon the limbic system, the broad architecture of which is shared with our evolutionary ancestors and mammalian cousins. However, what she shows, at best, is that these cognitive mechanisms arise relatively early in cognitive development. It remains possible that the physical and/or psychological architecture of our mature cognitive mechanisms replaces our primitive cognitive mechanisms. Moreover, even if the aetiology of our mature faculty for theoretical reasoning makes reference to infantile cognitive mechanisms, it does not follow that there are illuminating theoretical or conceptual connections between the two.

<sup>9</sup> Named after Elizabeth Spelke. See, for example, Spelke (1990).

<sup>10</sup> See Xu (1997: 378–9, 390). Earlier work by Bower (1974) suggested that infants are able to individuate objects in their environment from around five months. Xu (1997) also argues that one can profitably view ‘physical object’ as being a very general sortal expression, grasped by much younger infants. One might object, however, that sortals are intended to provide criteria



to recognise that an object cannot be both a ball and a bottle, or that a sapling and a mighty oak can be numerically identical.

There is also evidence that infants are able to individuate and classify objects according to their properties from around three months. Moreover, three-month-old infants appear to be aware of spatial relations such as aboveness and betweenness.<sup>11</sup> Slightly older infants, aged between seven and ten months, appear to individuate objects on the basis of conjunctions of properties.<sup>12</sup> Furthermore, they seem to notice dependencies between the properties and relations of one object and the properties and relations of another object. For example, infants aged ten months appear surprised when shown a video of atypical causal relations between billiard balls. A time lag between one billiard ball striking another, and the second billiard ball beginning to move, elicits greater than habitual attention.<sup>13</sup>

In short, the evidence from developmental psychology suggests that infants recognise in the actual world the elements of a KF-world: objects, properties, relations and dependencies.<sup>14</sup> In defence of (3), Maddy combines this finding with the earlier conclusion that KF-structuring is present in the physical world. The presence of KF-structuring in the physical world results in a strong evolutionary pressure in favour of logical cognition which is responsive to that structuring. Indeed, this type of logical cognition may be innate.<sup>15</sup> For, to paraphrase Quine, creatures inveterately wrong in their deductions have a pathetic but praiseworthy tendency to die before reproducing their kind.<sup>16</sup> Or, as Maddy remarks, ‘as long as KF-structuring is present, logic will serve us well’ (2007: 272). Experiencing the world in terms of Spelke objects exemplifying the

---

for counting objects of a particular kind and that the putative sortal ‘physical object’ fails in this respect. A demand that one counts the physical objects in a given area seems underspecified.

<sup>11</sup> Quinn (1994).

<sup>12</sup> Younger and Cohen (1983).

<sup>13</sup> Leslie and Keeble (1988).

<sup>14</sup> Notably absent, as Maddy allows, is an explanation of quantification in terms of our early cognitive capacities (2007: 262). This, she suggests, requires further empirical research.

<sup>15</sup> At least, many aspects of infants’ logical cognition are apparently pre-linguistic and universal abilities, shared with animals such as monkeys, pigeons and chicks (Maddy 2007: 268–9). The evidence concerning neonates is inconclusive, partly because of neonates’ limited visual sensitivity and consequent inability to respond to experimenters’ visual stimuli.

<sup>16</sup> See Quine (1969b: 126). See, however, Stich (1985, 1990) and Plantinga (2002) for reasons to doubt whether true beliefs enhance fitness.

properties, relations and dependencies of a KF-world, is at least a useful heuristic in this regard. It gives us largely reliable information from the outset, and helps us to navigate our early environment successfully. It is also worth noting that for the most part Spelke objects are portable or animate. Consequently, they are the sorts of objects which are most likely to relate to our basic survival needs.

We have, then, an interesting and provocative sketch for a naturalised philosophy of logic. In its favour, it seems correct that there should be some sort of correspondence between the structural features of logical cognition and the structural features of the world. For, if this were not the case, and if we assume a correspondence theory of truth, deductive validity would not necessarily preserve truth. Moreover, it is unsurprising that we are equipped from an early age with the cognitive tools for recognising logical structures in the world. We know that children have grasped the basic principles of logic from around the age of two, because they are beginning to use language by that point. It is unlikely that infants' logical reasoning capacities develop offline before that point, disconnected from experience and action. Aside from the aetiology, Maddy also offers us a controversial account of the truth conditions for logic. Although I shall return briefly to the topic of logic and physicalism, I do not intend to challenge either element directly. Rather I propose to look at what is absent from Maddy's account, namely an explanation of logic's prescriptive force.

#### *4.2 Justificatory and explanatory reasons*

For any given domain there are indefinitely many interesting and legitimate philosophical questions. It strikes me, however, that three types of questions are especially central to the philosophical enterprise – and, hence, to the enterprise of naturalised philosophy if it is to be a candidate successor discipline. We can ask whether the referents of the domain of discourse exist (and we can inquire into their mode of existence). We can ask about the relationships between our mental states and the objects within that domain, and in particular our cognitive access to them. And, we can ask about the value or normative status of objects within that domain. These questions roughly correspond to the sub-disciplines of

metaphysics, epistemology and ethics.<sup>17</sup> Moreover, assuming that the world does not harbour contradictions, the answers to the metaphysical, epistemological and ethical questions must be consistent. This requirement is widely recognised and enables intra-disciplinary criticism. For instance, moral realism is often criticised on the metaphysical grounds that it requires an extravagant and queer ontology.<sup>18</sup> And, mathematical platonism is often criticised on the epistemological grounds that we could have no cognitive access to causally inert objects.<sup>19</sup> Although it is perhaps less common to criticise metaphysical or epistemological theses on ethical grounds, there is no principled reason why we should not do so. For instance, Lewis' modal realism might be subject to the charge that it is incompatible with changes in the balance of good and evil in the pluriverse, and, so, forestalls the possibility of moral agency.

As I have sketched it above, Maddy's philosophy of logic is limited to addressing the first two types of question. This is not a criticism. If successful, an ontology and epistemology of logic would be achievement enough. Nevertheless, her account does need to be consistent with certain truths concerning the value and normative status of logic. For instance, it must be compatible with the fact that we are justified in reasoning logically. Alternatively expressed, it must be consistent with a vindicatory explanation of our logical practices. For, a non-vindicatory account would be self-defeating.<sup>20</sup> It is one thing to reason in a circle by relying on the logical principles that one then goes on to justify. Naturalists have no choice but to proceed thus, and can attempt to argue that the circle is a virtuous one. Indeed, given that rational thought and discourse presupposes logic, this appears to be a perfectly general predicament.<sup>21</sup> It is quite another thing, however, to rely on principles which one then undermines.

---

<sup>17</sup> I am not of course suggesting that this tripartite division is the only useful way to conceptualise the subject matter of philosophy. For instance, a further very good question is whether the objects are beautiful. Aesthetic value is importantly distinct from the sort of normativity under discussion insofar as its existence does not place obligations upon us – although it does of course give us reasons to act.

<sup>18</sup> I am of course thinking of Mackie's (1977) 'argument from queerness', which I shall discuss at length in chapter nine.

<sup>19</sup> Benacerraf (1973) is the modern *locus classicus* for this argument.

<sup>20</sup> See Williams (2000: 36–7) for more on vindicatory explanations.

<sup>21</sup> One manifestation of this predicament is the 'Cartesian Circle'. See further Frankfurt (1965).

A philosophy of logic, then, must provide an explanation of logic's prescriptive force, or be consistent with a complementary explanation. This explanation might be either intrinsic or extrinsic to logic. Employing the terminology of chapter one, either the truth-makers for logical propositions possess the requisite normative force to perform dual service as value-makers, or they connect up appropriately with an extra-logical source of justification. As we shall see below, Maddy is inclined towards the second option, justifying logical practice in terms of its instrumental relation to external, non-logical goals. For now, though, I propose to look at the first possibility and consider whether the above story might contain within itself a justification of our logical practices. That is, I would like to look at whether Maddy's naturalistic truth-makers for logic could serve as value-makers.

As claims (2) and (3) offer some appearance of a normative explanation, I would like to examine them in greater detail. In particular, I am interested in the type of reasons on offer. Although what I have to say concerns the nature of reasons in general, I shall focus on (2) for the sake of having a concrete example. Recall, then, that (2) is the claim that human beings believe logical truths because their most primitive cognitive mechanisms allow them to detect and represent the aforementioned features of the world. Now, this claim might be interpreted in at least two ways. The term 'because' invites a reason, but it is familiar that reasons can be explanatory or justificatory.<sup>22</sup> Whilst there is scope for debate concerning the best way to understand this distinction, I propose that we understand justificatory reasons as standing objectively in favour of an action or mental state.<sup>23</sup> By contrast, we can then understand explanatory reasons as merely allowing us to understand why someone did so-and-so, or believed such-and-such. Consequently, explanatory reasons do not necessarily have normative

---

<sup>22</sup> One sometimes encounters a similar distinction between normative and motivating reasons (e.g., Smith 1994, Schroeder 2007). I think this is potentially more problematic insofar as it can embody a commitment to motivational internalism. The idea behind the normative/motivating distinction is that we recognise a normative reason, which objectively stands in favour of an action, and that this recognition then constitutes a motivating reason. However, this contentiously assumes that external facts do not motivate directly.

<sup>23</sup> I do in fact favour externalism with respect to justificatory reasons. See Dancy (2000) for a full defence of this view. See also Parfit (2011: 31–8). In chapter eight, I defend the related view that reasons are not based on desires.

force or significance. If we learn that Nick deceived us because he is a compulsive liar, we gain an explanation but not a justification of his behaviour.

When Maddy argues that we believe logical truths because of our most primitive cognitive mechanisms, she is advancing a causal explanation. Although the relationship between the two is complicated, causal explanations are logically independent of justificatory reasons. It is true that discovering the causal explanation for a belief-forming mechanism can be relevant, in one sense, to whether one is justified in holding the resulting beliefs. For example, if one were to discover that one's confidence in the epistemic authority of a cult leader was the result of brainwashing, then the apparent warrant for the beliefs one had adopted on the basis of the leader's authority would be thereby undermined.<sup>24</sup> If one's reason for believing  $p$  was the leader's testimony then, upon discovering that one had been brainwashed, one would lose what one had taken to be the warrant for the belief that  $p$ . The reason for one's belief would have been debunked and, so, in the absence of further reasons, the rational course of action would be to withhold belief.<sup>25</sup> But, whilst this discovery bears on whether or not one ought to persist in the belief, the 'ought' in this case is subjective.<sup>26</sup> It is a matter of what one ought to believe given the available information. The question

---

<sup>24</sup> More commonly, it is difficult not to harbour the suspicion that one's moral or political beliefs might well have been different had one been brought up in a different environment. Similarly, one might uneasily recognise that commitments to philosophical theses tend to cluster geographically. For instance, G. A. Cohen (2000) observes that Oxford graduates of his generation were far more likely than their Harvard counterparts to accept the analytic/synthetic distinction. See White (2010) for further discussion.

<sup>25</sup> Whilst this would be the rational course of action, there is evidence of a widespread tendency towards belief perseverance in the face of disconfirming evidence (e.g. Jennings et. al. 1981). Harman (1984) plausibly attributes this phenomenon to the fact that people do not generally keep track of the justification for their beliefs once the belief is formed. This has the benefit of reducing mental clutter.

<sup>26</sup> Schematically, an objective ought has the form 'since  $x$  is the case,  $S$  ought to  $\phi$ ', whereas a subjective ought has the form 'since  $S$  believes  $x$  to be the case,  $S$  ought to  $\phi$ '. Thus, an objective ought is grounded in what is the case, whereas a subjective ought is grounded in what an agent believes to be the case. More colloquially, we might speak of 'there being a reason' as opposed to 'having a reason' respectively. The *locus classicus* for discussion of the distinction between objective and subjective oughts is Prichard (1932). Prichard asks us to consider whether the driver of a car ought to stop, or at least slow down, when approaching an intersection. Although it is natural to say 'yes', Prichard argues that from a certain objective standpoint 'there will be a duty to slow down only if in fact there is traffic' (1932: 93). The thought is that if a driver speeds through onto a main road which is in fact bare of oncoming traffic, no negative consequences will ensue. Hence, it will appear from a God's-eye perspective, or with retrospect, that there could have been no duty to slow down in the first place. More recently, Williams (1981: 102) raises a related set of concerns by considering whether one has reason to drink what one takes to be gin and tonic, but which is in fact petrol.

of whether or not one (objectively) ought to believe that  $p$  is independent of one's prior belief in the epistemic authority of the cult leader. It depends on the strength of the reasons for belief that  $p$ .<sup>27</sup>

There must be some set of explanatory reasons for the development of our capacity for logical thought and very likely some of these reasons will be in the vicinity of Maddy's suggestions. However, there is no guarantee that the explanatory reasons will be justificatory, or even relevant to the justification of our logical practices. In order to have normative force, the 'because' must be followed by a justificatory reason. However, it is unclear how the fact of our primitive cognitive mechanisms detecting and representing KF-structures could justify our logical practices. Consider, by way of analogy, a putative explanation of the widespread phenomenon of attributing greater weight to self-regarding reasons than to other-regarding reasons given in terms of the cognitive processes of vulnerable and dependent infants. Whatever the merits of this as an explanation, however deep human selfishness runs and however early it arises, it is no argument against ethical impartiality. This is not to deny that our cognitive development bears on ascriptions of praise and blame. We cannot appropriately blame someone for failures of thought or action entirely due to innate psychological limitations or a markedly impoverished environment. For, ought implies can and difficulty mitigates blame. In this way, facts concerning our cognitive development can serve as background conditions or enablers of justificatory reasons. Nevertheless, facts concerning our cognitive development do not determine the standards against which our thoughts and actions are measured. For, to repeat, explanatory and justificatory reasons belong to logically distinct categories.<sup>28</sup>

---

<sup>27</sup> This stands in contrast with epistemological internalists who hold that an agent's justification for her beliefs are necessarily cognitively available because internal. It seems to me, however, that internalism erroneously conflates subjective and objective oughts, or being justified in the beliefs one holds and holding justified beliefs. The latter, I take it, is the distinctively epistemological question because it is truth-orientated. Internalists pick up on something important and interesting by questioning whether someone is justified in holding one belief given her other beliefs. This question is particularly relevant to the epistemic evaluation of agents. But, to suppose that this exhausts the question of epistemic justification, or answers the question of whether the agent is justified in their belief *tout court*, is, I suggest, an error.

<sup>28</sup> Nagel illustrates this point with the following example: 'If ... someone says to me, "You only believe that  $2 + 2 = 4$  because you were in love with your second grade arithmetic teacher," that fails to count as a challenge. I may call up the long-buried image of Miss Gardbaum, with her

### 4.3 Bridging the Gap with Psychologism

One way in which one might attempt to bridge or collapse the gap between explanatory and justificatory reasons would be by endorsing logical psychologism. On this view, commonly attributed to Mill, the subject properties of logic are psychological.<sup>29</sup> Thus, logic concerns the laws of *actual* thought, as opposed to putatively independent laws governing thought. If this view were correct, then the prescriptive laws of logic would refer to descriptive psychological laws and we would at least know where to look to justify our logical reasoning practices.<sup>30</sup> Logical norms would either be relative to the particularities of an individual's psychology, or, more plausibly, grounded in patterns of reasoning common amongst a population. The existence of common reasoning practices might then be explained in terms of Maddy's proposed primitive cognitive mechanisms or perhaps an innate logic module.<sup>31</sup>

But, whilst logical psychologism is a possible move at this point in the dialectic, it is a highly unfashionable view.<sup>32</sup> This is largely due to a number of Fregean criticisms, which have persuaded most philosophers that the logical and the psychological are fundamentally incompatible.<sup>33</sup> For instance, Frege argued that whereas logical laws are universal, certain and precise, psychological laws

---

soft hair, prominent bosom, and dark blue skirt powdered with chalk dust, and acknowledge that yes, I was in love with her and wanted to believe everything she told me – but these reflections will be powerless to make me reconsider my conviction that  $2 + 2 = 4$ , because it lies beyond their reach and does not depend on anything which they call into question.' (1997: 56)

<sup>29</sup> The appropriateness of this association is denied by some Mill scholars, perhaps most notably by Skorupski (1989). Godden (2005: 115–7) provides a clear overview of the debate. In all likelihood, Mill was not fully consistent on the issue.

<sup>30</sup> In the broader context of naturalised epistemology, Kornblith characterises psychologism in similar normative terms as 'the view that the processes by which we ought to arrive at our beliefs are the processes by which we do arrive at our beliefs' (1994: 8). Such a perfect coincidence may appear vanishingly unlikely, unless as the result of stipulative definition. However, Kornblith suggests that the prospects are much improved by evolutionary considerations concerning the fitness-enhancing qualities of true beliefs.

<sup>31</sup> See, for instance, Rips (1994). Hanna (2006: 46–52) also argues that we have an innate capacity for logic, partly on the basis that this is presupposed by Chomsky's theory of universal grammar.

<sup>32</sup> As Aach (1990: 315) bluntly puts it 'psychologism is the view that logic can be explicated by psychology, and though it was widely accepted during the last century, today it is considered dead'. See further Jacquette (1997) on the historical development of the negative connotations attaching to the term 'psychologism' – connotations which he then suggests are not entirely fair. See also Pelletier, Elio and Hanson (2008) for an extensive survey article.

<sup>33</sup> Husserl's *Prolegomena to Pure Logic* is also important in the history of anti-psychologism. See Kusch (1995: 203–10), for a discussion of the relative significance of Frege's and Husserl's anti-psychologistic in the early decades of the twentieth century.

are particular, merely probable and imprecise. Perhaps most influential, however, has been Frege's argument that psychologism conflates being true with being taken to be true.<sup>34</sup> In this context, he offered the following attractive metaphor:

If being true is thus independent of being recognized as true by anyone, then the laws of truth are not psychological laws, but boundary stones set in an eternal foundation, which our thought can overflow but not dislodge. (1893: xvi)

Frege meant of course to affirm the antecedent. The idea is that whereas descriptive psychology is confined to studying the conditions under which we take inferences to be valid, logic concerns the conditions under which inferences are in fact valid. Consequently, to reduce the logical to the psychological is to conflate the objective with the subjective. This is a disturbing prospect given that logic is widely held to be paradigmatically objective. Subjectivism with respect to logic appears just a short hop from global epistemological relativism, and an attendant pernicious scepticism.

In defence, however, an advocate of psychologism can point out that even if the content of logic depends on subjective mental states, the facts concerning those mental states need be no less objective than other scientific facts. Moreover, he might argue, the fact that reality is mind-independent does not entail that being true is mind-independent. For, he might continue, being true is a property of beliefs, not a property of a mind-independent reality. It is a substantive further question whether truth is objective or, as it sometimes put, verification-transcendent. Here, the advocate of psychologism might find support in the kind of anti-realist considerations which motivated Brouwer's intuitionistic logic and Dummett's identification of truth with warranted assertability. Although this raises questions far beyond the scope of the present discussion, it does at least seem that psychologism has more resources than is commonly appreciated in relation to the charge of subjectivism.

---

<sup>34</sup> The same objection can be traced back at least to Kant. See further Jacquette (2001: 262–6).



On the subject of normativity, Frege criticises psychologism on the grounds that it conflates justificatory reasons with explanatory reasons:

With the psychological conception of logic we lose the distinction between the grounds that justify a conviction and the causes that actually produce it. This means that a justification in the proper sense is not possible. (1897: 147)

This is essentially the same type of argument I have been trying to make. We should, however, note that collapsing this distinction between justification and cause is precisely what might motivate logical psychologism as a way of accounting for the normativity of logic. To criticise logical psychologism simply on the grounds that it conflates justification and cause is therefore to beg the question. Although psychologism removes any further question concerning the justification of our logical convictions in terms of correspondence with, or the prescriptive force of, non-psychological logical content, two possibilities remain. On the one hand, it is compatible with our logical convictions being justified by non-logical facts. I shall come to this possibility shortly, and discuss it at length in the following chapter. On the other hand, it is possible that the causal reasons coincide with the justificatory reasons. The alleged conflation then becomes an identity claim. To head off this possibility we need to show, as I attempted above, that it is a distinction with a difference and that justificatory reasons are required for explaining the normative force of logical inference independently of the causal or explanatory reasons. This is what I understand Frege to mean by the qualification ‘justification *in the proper sense*’.

Although I think that psychologism is more defensible than commonly assumed, I am not aware of any contemporary philosopher who is prepared to endorse it.<sup>35</sup> Maddy is certainly in the mainstream insofar as she is persuaded against psychologism by Frege’s arguments (2007: 200–1). Whilst she bases her account of logical cognition on developmental psychology, she maintains that

---

<sup>35</sup> The following are near misses. Dale Jacquette (1997; 2001) defends psychologism without going so far as explicitly endorsing it. Robert Hanna’s (2006) ‘logical cognitivism’ is psychologistic insofar as it takes logic to be constructed by a logic faculty in the mind. Hanna denies, however, that logic is explanatorily reducible to empirical psychology (2006: 27). Pascal Engel (2005) proposes a psychologistic account of logical reasons, but not of logical content. It is also natural to read Quine’s (1969a) essay ‘Epistemology Naturalized’ as a pro-psychologism rallying cry. See, however, Margolis (1997: 300–6) for an alternative reading.

logic is made true by structural features of the physical world.<sup>36</sup> Nevertheless, psychologism and physicalism have important features in common, allowing us to carry forward and generalise some aspects of our discussion. Notably, both positions are attempts to provide truth conditions for logic which are contingent, *a posteriori* and fit into the causal structure of the world as described by the natural sciences. These common features make psychological or physical facts the obvious choices for the content of a naturalised philosophy of logic. However, it is difficult to see how either category of facts could function as value-makers in addition to truth-makers. As argued above, they do not do so just in virtue of their causal/historical relationship to our logical reasoning capacities. For, when we look solely at the fact that we have come to reason in a particular way due to a particular chain of events, it is perfectly compatible with it being the case that we ought not to reason thus. This is because there is no necessary connection between explanatory reasons and justification.

It is also difficult to see how physical or psychological states could function as value-makers on the basis of intrinsic properties that are cognitively available from a naturalistic, objectifying perspective. Here, we might recall the Humean considerations described in chapter one. The recognition of, say, K-F-structuring from an impersonal, scientific perspective does not have normative authority over us and does not oblige us to reason in particular ways. Given our recognition that the physical world is structured in a particular way, or that people generally reason in accordance with particular logical norms, it remains an open question whether one ought to reason in this way or that. But, if this is especially clear in the case of contingent physical facts and mental states, the issue does not essentially depend on the naturalists' preferred ontology of logic. Indeed, it is common ground that a strictly scientific perspective does not recognise the intrinsic magnetism of *sui generis* norms. Like the moral property

---

<sup>36</sup> Of course, physicalism with respect to the content of logic comes with its own well-known set of difficulties. For instance, if K-F structures are to have the right modal properties to function as truth-makers for logical propositions, then logical truths had better be contingent. If the physical world could have been different, as is apparently the case, then so could the content of logic. Maddy accepts these revisionary consequences and attributes the widespread conviction that logic is necessarily true to 'psychological realities' rather than any deeper metaphysical truth (2007: 273). See also Bigelow (1988) for a defence of physicalism in the philosophy of mathematics and consideration of parallel debates.

of to-be-pursuedness, the property of to-be-thoughtness is queer and alien to the scientific world-view with its focus on causal and constitutive relations. I shall offer some explanation of this scientific alienation in chapters six and nine when I discuss the nature of practices and the role of the emotions in normative cognition respectively. As there is general agreement, however, I shall not dwell further on the issue here.

Whilst I have already argued that there is no necessary covariance between explanatory reasons and justifications, it remains possible that they coincide in particular cases. Even if naturalistically construed truth conditions for logic are not intrinsically prescriptive, a naturalist might argue that as a matter of fact the explanation and justification of our logical beliefs coincide. Given that we are dealing with normative reasons, it is difficult to find a non-contentious example of coincidence between explanation and justification. In general, however, if we have a theory which makes success contingent upon something being done in the right sort of way, or as the result of the right sort of process, then there is the possibility that explanations and justifications will coincide when the success conditions are met. For instance, one plausible analysis of knowledge is as true belief acquired through the exercise of intellectual virtue.<sup>37</sup> If this analysis is correct, intellectual virtue can both explain and justify belief. For instance, my belief that there is a plant on my desk is both explained and justified by the skilful exercise of my perceptual capacities. Might it not be, then, that certain psychological or physical explanations similarly justify *and* explain our practices of logical reasoning? Well, for all I have said so far, this could indeed be the case. However, as I shall now argue, these examples of coincidence between explanation and justification require a further source of justification.

In order to see that these examples of coincidence depend on an independent normative standard, consider a case in which I come to believe that *p* through the exercise of intellectual virtue, but in which, due to sheer bad luck, my belief that *p* is false. Perhaps I have been tricked by an evil demon or dropped into

---

<sup>37</sup> See, for instance, Zagzebski (1996) and Sosa (2007). I defend a version of virtue epistemology in chapter seven.

Goldman's notorious Barn County. Is my belief that  $p$  justified in this scenario? Clearly, given that the belief is the product of intellectual virtue, there is a sense in which I am justified in holding the false belief. I am not blameworthy for being the victim of epistemic bad luck and, indeed, I subjectively ought to believe that  $p$ . However, the fact that I am justified in holding the belief does not entail that the belief is justified, unless one fails to distinguish between subjective and objective oughts.

As  $p$  is false, we know that it is neither justified in virtue of its correspondence with the way things stand in the world, nor logically entailed by the content of some true foundational belief. Perhaps it will be claimed that  $p$  is justified because it coheres with (some subset of) my other beliefs, the conjunction of which we might call  $q$ . There are of course a number of ways in which we might understand the coherence relation. The approach does not, however, look promising. If, for example, we understand coherence in a strong sense according to which a proposition coheres with a set of propositions just in case it is entailed by that set, then clearly  $p$  cannot receive any justification from  $q$ . For, if  $q$  is false then  $q$  entails both  $p$  and not- $p$ . And, if  $q$  is true then it entails not- $p$  (because  $p$  is false). Even if we understand coherence in a weaker sense, such as, say, mutual explanatory support, it does not seem to provide any justification. For, it seems plausible that any explanatory support for  $p$  must proceed through the truth of  $q$ . If  $q$  were false it could hardly explain or justify  $p$ . A false proposition entails everything and so does not explain anything in particular. But if the falsehood of  $q$  cannot justify  $p$  then neither can the truth of  $q$ . For, coherence is a symmetrical relation. So, if the truth of  $q$  provides explanatory support for the falsehood of  $p$ , then the falsehood of  $p$  must provide explanatory support for the truth of  $q$ . But we have already denied that the explanatory support or justification for one proposition can come through the falsehood of another.

If my belief that  $p$  is not justified by the way matters actually stand, or by its coherence with the content of my other beliefs, perhaps it is justified simply because it is the product of intellectual virtue. For, certainly, I have suggested that in cases of the successful exercise of intellectual virtue the justification and

explanation of a belief coincide. However, an intellectual virtue only counts as virtuous because of its relationship to some epistemically valuable end such as truth, knowledge or what we might generally call cognitive gain. If some disposition were not reliably conducive to cognitive gain, then quite clearly it would not be an intellectual virtue. The disposition might be beautiful, or moral, or otherwise valuable, but intellectual virtues are identified as such from a truth-orientated, epistemic perspective. Thus, justification is transmitted from the epistemic good of cognitive gain to intellectual virtue. Intellectual virtue does not constitute an independent source of epistemic justification in addition to the justification transmitted from the goal of cognitive gain. That would be a case of double counting, for a stable disposition would not be a virtue at all if not appropriately related to the right end. Consequently, the fact that a belief is acquired through the exercise of intellectual virtue does not guarantee that the belief is justified. Certainly, it does nothing to justify the content of the belief, which can only be justified by objective reasons standing in its favour. So, intellectual virtue will only serve naturalised logic as a model of coincidence between explanation and justification if there is an independent source of justification linking them together.

We have seen that the naturalised truth conditions Maddy proposes for logic are not intrinsically normative and that the psychological explanation she gives us of our capacity for logical thought is insufficient to justify our logical beliefs. We have also seen, however, that explanations and justifications can be linked together by an independent source of normativity. Although Maddy does not provide an explicit account of the normative force of logical inference, we can reasonably anticipate the approach she would take from her discussion of another gap in her model.

There is some distance to travel between rudimentary logic based on K-F structures and full-blown classical logic. Although objects, properties, relations and certain dependencies are present in infant cognition, we do not as yet have our full complement of connectives and relations of implication. Nevertheless, Maddy claims that we are entitled to introduce such features:

For the purposes of logic, of course, it is a boon, bringing the last of the basic connectives into the truth-functional fold and supporting the full range of classical tautologies, including the interdefinability of the connectives. (2007: 287)

Again:

The justification must be, as always, that they make it possible to achieve results that would otherwise be impossible or impractical, and that they do so without introducing any relevant distortions. (2007: 288)

Thus, the justification for moving from rudimentary logic to classical logic depends on normative considerations, and assumes that the normativity involved is instrumental.<sup>38</sup> Accepting certain principles is a ‘boon’ because it enables us to achieve certain desired results. From context, the word ‘relevant’ in the second of the above quotations means something like ‘relevant relative to our background concerns’.

Maddy does not develop this instrumentalist approach to normativity or attempt to supplement her naturalised account of logic with it. Nevertheless, it is a popular strategy and it is easy to see how it could be employed to yoke a certain sort of justificatory reason to the explanation Maddy offers for our beliefs concerning logic. Our logical convictions, she might suggest, are both explained and justified by their instrumental relation to the role true beliefs play in satisfying our practical and theoretical interests. Rather than evaluating this possibility immediately, however, I propose to shift into the territory of naturalised epistemology in order to focus on the more fully developed accounts of instrumental normativity offered by Stich, Kornblith and Papineau amongst others.

---

<sup>38</sup> Maddy explicitly proposes an instrumentalist approach to axiom justification in her work on mathematical naturalism. She suggests that we should ‘frame a defense or critique of a given [axiom candidate] in two parts: first, identify a goal (or goals) of [set-theoretic] practice, and, second, argue that [adopting the axiom] in question either is or is not an effective means towards that goal’ (1997: 97).

## **Chapter Five** **Naturalised Epistemology and Instrumentalism**

W. V. Quine exerted perhaps the strongest influence on the ‘naturalistic turn’ of the mid to late twentieth century. A central element of his programme was an insistence on the possibility and desirability of naturalising epistemology.<sup>1</sup> This chapter looks at whether naturalised epistemology can account for epistemic normativity. In the previous chapter I considered the prospects for identifying naturalistically acceptable explanatory reasons with normatively significant justificatory reasons. In this chapter I propose to examine the strategy of reducing epistemic normativity to instrumental normativity.<sup>2</sup>

In Section 5.1, I look at the relationship between naturalised epistemology and a traditional understanding of the nature of epistemology. I also argue that epistemology is an essentially normative discipline. In section 5.2, I describe Quine’s proposal that we reduce epistemic normativity to instrumental normativity, and the more fully developed accounts of Stich (1990), Kornblith (2002) and Papineau (2003). In section 5.3, I argue that we are in fact subject to epistemic norms regardless of instrumental considerations. In part, I defend this view by looking at the case of trivial knowledge. Finally, in section 5.4, I argue that instrumentalists are unable to justify their appeals to truth and consequently unable to explain the distinctively epistemic dimension of epistemic normativity. As a result, instrumentalism does not provide an adequate account of epistemic normativity, and, so does not provide support for naturalism.

### *5.1 Naturalised Epistemology and Normativity*

In his seminal essay on the subject, Quine described the programme of naturalising epistemology as follows:

Epistemology, or something like it, simply falls into place as a chapter of psychology and hence of natural science. It studies a natural phenomenon, viz., a physical human subject. The human subject is accorded a certain experimentally controlled input – certain patterns of irradiation in assorted

---

<sup>1</sup> I borrow the phrase ‘naturalistic turn’ from P. M. S. Hacker (2006: 231).

<sup>2</sup> See, for example, Brown (1988), Giere (1989), Laudan (1990), Maffie (1990), Papineau (2003a), Solomon (2001), Hubin (2001), Kornblith (2002), Janvid (2004) and Huss (2009). Berger (2003: 372) prefers the label ‘if-thenism’.

frequencies, for instance, and in the fullness of time delivers as output a description of the three-dimensional world and its history. (1969a: 82–3)

In a later essay, he wrote as follows:

Naturalism does not repudiate epistemology, but assimilates it to empirical psychology. Science itself tells us that our information about the world is limited to irritations on our surfaces, and then the epistemological question is in turn a question within science: the question how we human animals can have managed to arrive at science from such limited information. (1981: 72)<sup>3</sup>

So, Quine saw the central task of epistemology as understanding the causal relations between the evidence of our senses and our mental representation of the world. He argued that this relationship between the ‘meagre input and the torrential output’ (Quine 1969a: 83) should be investigated wholly within the remit of empirical psychology. Thus, the philosophical problems associated with the theory of knowledge should be tackled using empirical methods, and a constrained epistemology should be subsumed by the natural sciences. Metaphysics or first philosophy should no longer be pursued, and distinctly philosophical methods should be consigned to the unenlightened past.

In short, Quine proposed that we replace epistemology, or its rump, with empirical psychology. However, as I argued in chapter three, if one discipline is to replace another then there must be substantial continuity between the original discipline and its successor. Thus, it would be a serious objection to Quine’s proposal if psychology and epistemology dealt with largely different questions. Although Quine’s proposal was deliberately revisionary, he still needed to show that psychology answers questions drawn from the core of traditional epistemology. For, otherwise, his claim that we should pursue psychology because traditional epistemology has failed would be a *non sequitur*. Recognising this requirement, he wrote:

---

<sup>3</sup> Quine similarly claims that it is a scientific finding that ‘information about the world reaches us only by forces impinging on our nerve endings’ (1981: 181). Although this is not the place to pursue the matter, I seriously doubt that Quine is entitled to these claims. If, however, he were so entitled, the simple expedient of providing references to the relevant scientific research would have established his point.



A far cry, this, from the old epistemology. Yet it is no gratuitous change of subject matter, but an enlightened persistence rather in the original epistemological problem. (1974: 3)

When Quine referred to the ‘original epistemological question’, he had in mind the following: ‘Given only the evidence of our senses, how do we arrive at our theory of the world?’ (1974: 1). This, he suggested, was the question pursued by Berkeley, Locke and Hume, and the question which should once more become the focal point of epistemology. But, setting aside Quine’s rather suspect, broad-brush view of early modern empiricism, it is surprising to encounter the proposal that epistemology can be reduced to a single question. In his last book, William Alston suggested that ‘a broad conception of epistemology would be *philosophical reflection on the cognitive aspects of human life*’ (2005: 3). This seems closer to the mark. Since antiquity, epistemologists have been interested in understanding the nature of knowledge and related cognitive states quite generally. Under what conditions are we in a state of knowledge? Does knowledge require certain foundations, or is coherence among a set of beliefs sufficient? How does knowledge differ from mere (true) belief or opinion? Is it possible to provide an analysis of knowledge? Is certainty possible and, if so, in which fields of enquiry?

Most importantly for present purposes, epistemologists have been centrally concerned with normative questions.<sup>4</sup> Under what conditions are we justified in believing some proposition? What, in any case, does epistemic justification primarily bear upon? Acts? Beliefs? Agents? Are the facts which justify beliefs internal or external to the knower’s mind? Why is knowledge distinctively valuable as opposed to, say, true belief?<sup>5</sup> Beyond these foundational concerns, applied questions also arise. For, once the conditions for being in a state of knowledge are understood, and given that knowledge is a valuable cognitive state, the epistemologist may find himself in a position to recommend certain strategies of belief formation and caution against others. He

---

<sup>4</sup> Whilst normative questions have always occupied epistemologists, it is possible to discern a recent heightening of interest, possibly associated with the renaissance of virtue epistemology. In an essay of the same name, Riggs (2008) has labelled this trend ‘the value turn in epistemology’.

<sup>5</sup> This has recently become known as ‘the value problem’ in epistemology. See Pritchard (2007) for an overview.

may, for instance, have advice to offer concerning the social dynamics of belief formation. Or, he may be alert to instances of epistemic injustice.<sup>6</sup>

There is, then, a widespread sentiment that epistemology has legitimate normative concerns. Moreover, there are good reasons to suppose that this sentiment is warranted. For instance, we might notice that normative concepts are ubiquitous features of epistemic discourse. We make statements such as ‘you ought to apportion your beliefs to the evidence’. We call beliefs ‘appropriate’, ‘fitting’ and ‘reasonable’, as well as making the converse negative judgments. We use normative vocabulary to describe an agent’s epistemic character. One can, for example, be ‘honest’, ‘responsible’, ‘courageous’ or ‘negligent’ in pursuing one’s inquiries.<sup>7</sup> Even Quine refers to a type of epistemic ‘duty’:

The purpose of concepts and of language is efficacy in communication and in prediction. Such is the ultimate duty of language, science, and philosophy, and it is in relation to that duty that a conceptual scheme has finally to be appraised. (1980: 79)

Even if the wording is slight carelessness on Quine’s part, the temptation to employ the language of duty is revealing.<sup>8</sup>

We might also take into consideration a line of argument suggested by Jaegwon Kim, which goes some way to explaining why epistemology is a normative concern. Taking as his starting point the classical tripartite conception of knowledge as justified true belief Kim observes that justification is the only distinctively *epistemic* component.<sup>9</sup> As he notes ‘neither belief nor truth is a

---

<sup>6</sup> See, for example, Fricker (2007).

<sup>7</sup> See Haack (1997: 30) for further discussion of normative language and epistemic virtue.

<sup>8</sup> Morton White (1986: 652) draws our attention to this passage as an illegitimate (for Quine) use of deontological vocabulary. In response, Quine tells us that when he referred to duty he ‘was using the word somewhat as when we speak of a heavy-duty cable or tractor. It was what language, science, and philosophy are for, as eyes are for seeing’ (1986: 665). I shall consider the teleological aspect of Quine’s position in the next section.

<sup>9</sup> Or, if not ‘justification’, then some close normative relative. Plantinga, for example, holds that the difference between knowledge and mere true belief is a matter of warrant. He also emphasises the normative aspect of epistemology: ‘To say that a belief is *warranted* or *justified* for a person is to evaluate it or him (or both) *positively*, his holding that belief in his circumstances is *right*, or *proper*, or *acceptable*, or *approvable*, or *up to standard*.’ (Plantinga

specifically epistemic notion: belief is a psychological concept and truth a semantical-metaphysical one' (1988: 383). It is true that Gettier cases have convinced most philosophers that knowledge is not exactly justified true belief. Nevertheless, there is good reason to suppose that whatever it is that makes the difference between knowledge and mere true belief is normative (or that knowledge is an essentially normative mental state). For the possession of knowledge brings credit to an agent in a way that the possession of mere true belief does not.<sup>10</sup> Moreover, shifting the focus of evaluation from the agent to his mental states, knowledge appears to confer added value upon true beliefs.<sup>11</sup> As Kornblith observes, 'to say that a belief is an item of knowledge is to praise it in a certain way; it is to approve of it as meeting our cognitive ideals; it is to recommend it' (2002: 159). Hence, the distinctively epistemic component of knowledge must be normative insofar as being in a state of knowledge is subject to positive epistemic appraisal in a way that mere possession of a true belief is not.

I am emphasising the fact that epistemology is a normative discipline because it has struck a number of philosophers that the normative is absent from Quine's account.<sup>12</sup> Moreover, it is not surprising that they have thought this. Scientific theory, Quine tells us, 'stands proudly and notoriously aloof from value judgements' (1973: 49). Consequently, it might appear that naturalised epistemology must stand similarly aloof if it is to become 'a chapter of psychology and hence of natural science' (1969a: 82). As Larry Laudan writes:

---

1993: 3). Although Sartwell (1992) has argued that knowledge is simply identical to true belief, he does not deny that an account of what justifies knowledge, rather than true belief, is both necessary and important to epistemology. In this context, see also Goldman (1999: 23–6) for discussion of (mere) true belief as 'weak knowledge' or 'W-knowledge'.

<sup>10</sup> The so-called 'credit thesis' is widely accepted amongst epistemologists. For a full defence see Greco (2003). See, however, Lackey (2007, 2009) for a dissenting voice.

<sup>11</sup> I return to this idea later in the chapter. However, The full explanation will have to wait until section 7.3 below.

<sup>12</sup> See, for example, Putnam (1982: 19) and Kim (1988: 388–9).

The likes of Descartes, Leibniz and Kant were keen to say how we *ought* to form our beliefs and how we *should* go about testing our claims about the world. Science, by contrast, does not appear to traffic in such normative injunctions; it describes and explains the world but it does not preach about it. (1990: 45)

*Pace* Quine, Laudan might have added Berkeley, Locke and Hume to his list of preachers. The concern is that naturalised epistemology can only be descriptive, explaining how, as a matter of fact, we acquire beliefs. If naturalised epistemology is to count as epistemology at all, it must address this concern. In the next section, I look at how this might be done.

### *5.2 Instrumentalist Approaches to Epistemic Normativity*

Even if the normative tasks of epistemology are displaced from the foreground in its naturalised variant, some account is required. For, as we have noted, naturalised epistemology cannot replace traditional epistemology unless it addresses the same subject matter. And, as I have just argued, traditional epistemology has normative concerns at or near its core. Consequently, despite Quine's general indifference towards normative philosophy, he does make a natural and influential suggestion. It requires that we bring into focus the end or *telos* of epistemology:

Naturalization of epistemology does not jettison the normative and settle for the indiscriminate description of ongoing procedures. For me normative epistemology is a branch of engineering. It is the technology of truth-seeking, or, in a more cautiously epistemological term, prediction. Like any technology it makes free use of whatever scientific findings may suit its purpose ... There is no question here of ultimate value, as in morals; it is a matter of efficacy for an ulterior end. The normative here, as elsewhere in engineering, becomes descriptive when the terminal parameter is expressed. (1986: 664–5)

Quine seeks to keep science aloof from value judgements by denying that there are *sui generis* norms *within* science. There are no intrinsically normative epistemic properties which it is the purpose of a naturalised epistemology to disclose. Rather, scientific values are contingent upon the aims of science, and apparently normative discourse turns out to be merely descriptive.

For Quine, identifying the ‘terminal parameter’ is a relatively straightforward empirical matter. One must find out what it is that scientists aim at in their practice. The answer Quine gives is that the primary scientific aim is to ‘maximise prediction’ or, to develop a theory ‘that will anticipate as many observations as possible, getting none of them wrong’ (1973: 137).<sup>13</sup> In line with the conclusions of chapter three, a better characterisation of the aims of science might be the production of knowledge in accordance with the changing norms of professional scientific journals and the regulative ideal of subjectless knowledge. In any case, questions of justification are reduced to questions concerning the reliability of belief-forming mechanisms with respect to scientific aims, and these can be answered empirically and independently of any particular stance or perspective. There is no further distinctively philosophical work to be done. As Alcázar (1993: 315) describes it, one ‘can happily give normative advice since this is the result of combining two elements which are not normative at all’. Epistemic rationality turns out to be a variety of instrumental rationality, or the rationality of taking the appropriate means to one’s ends.

Before looking in more detail at particular instrumentalist proposals, a few comments on the nature of instrumental value might be helpful. In particular, we might note that instrumental value is not the opposite of intrinsic value as some instrumentalists seem to think.<sup>14</sup> The appropriate contrast is between intrinsic and extrinsic value on the one hand, and instrumental and final value on the other hand.<sup>15</sup> The instrumental/final distinction distinguishes between two ways in

---

<sup>13</sup> Laudan (1990: 48–9) advocates an instrumentalist position, but maintains that the aims of science, and therefore epistemic norms, change over time as scientific practice changes. Although I set aside this complication in the following discussion, it does seem that Laudan’s ‘historicism’ is the appropriate position to take given a naturalistic denial of extra-scientific norms. Scientific practice is historically contingent and so must be its norms, within certain continuity constraints.

<sup>14</sup> See, for instance, Kornblith (2002: 161) and Stich (1990: 131).

<sup>15</sup> See Korsgaard (1983) for detailed discussion of these distinctions.

which something might be valued, either as a means or as an end. This type of value can only arise in relation to some goal. By contrast, the intrinsic/extrinsic distinction distinguishes two ways in which something might be valuable, respectively in virtue of its non-relational or its relational properties. These distinctions cut across one another. For instance, a gun may have instrumental value as a means to a murder. However, this does not entail that the gun has extrinsic value. The murder may be an end entirely without value. So, the gun, whilst instrumentally valuable, can hardly be extrinsically valuable in virtue of its relation to the murder.<sup>16</sup> Although it is true that instrumental value *can* be a species of extrinsic value, this is not always the case.<sup>17</sup>

Let us return from our brief excursus into value theory and consider how instrumentalism with respect to epistemic normativity might be developed. Although Quine advocated instrumentalism with little elaboration or defence, it has struck many naturalists as a promising approach.<sup>18</sup> As we saw at the end of the previous chapter, Maddy relies on instrumentalism as she fills out her naturalised account of logic. However, the question arises, ‘how and why does an epistemic end oblige us?’ If naturalists can escape the charge of eliminating normativity by appealing to instrumental norms contingent upon the aims of scientific practice, they still require an account of the normative force of those aims. If scientific practice is directed towards the production of particular types of true propositions, whence does that end derive its normative force?<sup>19</sup> As Hilary Kornblith observes:

We cannot rest content with Quine’s seemingly innocent suggestion that epistemic norms ‘become descriptive when the terminal parameter is expressed’, for we need to know what the source of this terminal parameter is. What, ultimately, is the source of epistemic normativity? (2002: 139)

---

<sup>16</sup> It might be suggested that something with instrumental value must always have extrinsic value insofar as it bears a relation to the satisfaction of an agent’s desires. In response, I would deny that there is value to the satisfactions of an agent’s desires *per se*, unless the desire is for something we have reason to value, i.e. something which is valuable. I shall return to this topic in chapter eight when I consider the relationship between reason and desire.

<sup>17</sup> Riggs (2008: 315) is mistaken in thinking that instrumental value is always a type of extrinsic value.

<sup>18</sup> See footnote 2 above.

<sup>19</sup> As Nietzsche asked in the opening section of *Beyond Good and Evil*, ‘Granted we want truth: why not rather untruth? See Williams (2002: 13–19) for a very good account of Nietzsche’s position on the value of truth.

I shall look at Kornblith's answer a little later. Although I shall argue that his answer is unsatisfactory, the question is a good one.

The important point, then, is that normative reasons cannot be endlessly transmitted without some genuinely normative origin.<sup>20</sup> A natural way in which we might try to locate the source of epistemic normativity is by considering the value of true beliefs.<sup>21</sup> If true beliefs are instrumentally valuable, we have a certain sort of reason to acquire them. Moreover, it is highly plausible that true beliefs are instrumentally valuable, for it seems likely that without an accurate map of our environment we would fare very badly. The thought then emerges that we should adopt particular patterns of inference because they are conducive to forming true beliefs, which, in turn are conducive to satisfying our goals.

I shall look directly at truth-orientated accounts of epistemic normativity shortly. By way of an approach to this topic, however, I want to look at a form of instrumentalism which does not treat truth as the end of epistemic activity. In particular, Stephen Stich (1990) has argued against the plausible claim that the possession of true beliefs has pragmatic value. Indeed, he has claimed that on becoming clear about the nature of truth, one sees that true beliefs have no value whatsoever. Essentially, he argues that there are many truth-like relations, and no good reason to prefer one to another other than their propensity to enable us to satisfy our goals. The details of his arguments are not entirely germane to our theme, but the relevant upshot is Stich's recommendation that we replace the idea of truth as the terminal parameter of enquiry, with other valuable ends. He writes:

In evaluating systems of cognitive processes, the system to be preferred is the one that would be most likely to achieve those things that are intrinsically valued by the person whose interests are relevant to the purpose of evaluation. So, for example, if the issue at hand is the evaluation of Smith's system of cognitive processes in comparison with some actual or hypothetical alternative, the system that comes out higher

---

<sup>20</sup> Cp. Aristotle's *Nicomachean Ethics* 1094a18-23.

<sup>21</sup> One might want to deny that beliefs are strictly true or false. It might be more accurate then to say 'the value of possessing beliefs with true propositional content' than the 'value of true beliefs'.

on the pragmatist account of cognitive evaluation is the one that is most likely to lead to the things that Smith finds intrinsically valuable. (1990: 131)

Now, we might interpret this in two ways: according to its letter or according to its spirit. Thus, by ‘the things that Smith finds intrinsically valuable’, we might understand either ‘the things that Smith finds to be valuable in virtue of their non-relational properties’ or ‘the things that Smith finds valuable regardless of whether they promote some *further* end of his’. As it is hard to see how the intrinsic/extrinsic distinction would be relevant to Stich’s position, I suspect he intended the latter. But, if this is correct, Stich has conflated intrinsic value with final value as cautioned against above. Although this could be dismissed as no more than carelessness, it may also be the result of adhering to a position that ties value to the satisfaction of personal goals. Otherwise, it is hard to see why the very different distinctions would be run together.

In any case, Stich’s account has several more worrying features. We might notice, for example, that it is a relativistic account of epistemic normativity in the sense that two individuals ought to have different beliefs depending on their respective systems of values. This, however, is incompatible with the highly plausible thought that epistemic justification is universalizable. As Ernest Sosa writes, ‘if a belief is epistemically justified, then any belief similar to that belief in all relevant respects would be equally epistemically justified’ (1993: 50). If Smith and Jones both hold the same belief on the same grounds then they are equally justified in holding it, whatever its relation to the achievement of their goals. If it were suggested that Smith’s, but not Jones’, belief is epistemically justified, then it seems to me that we would have lost our grip on the notion of *epistemic* justification altogether. Better by far to say that epistemic justification and pragmatic utility can come apart. Moreover, just as epistemic justification is constant between persons, so it is independent of the fact that an individual’s goals can vary over time. A belief which once frustrated my goals may later further them, but the belief does not thereby enjoy epistemic justification which it did not previously have. For, otherwise, one could justify one’s beliefs by working on one’s desires.



One manifestation of the failure to account for the universalizability of epistemic reasons is that Stich's account does not seem to allow for the right account of what psychologists call 'motivated' or 'hot' reasoning.<sup>22</sup> Hot reasoning occurs when lines of thought are promoted on the basis of emotional salience, or when cognition is directed towards non-cognitive ends. It is true that all conscious human cognition is hot in the sense that it is guided or influenced by feelings such as certainty.<sup>23</sup> However, cognition *directed* at non-cognitive ends is widely derided as 'sham' reasoning.<sup>24</sup> Our practices of epistemic evaluation indicate that we evaluate a subject's beliefs regardless of what we think of her motivation or desires, or whether indeed we know anything about them. At least, this is the regulative ideal we are guided by when, for instance, we evaluate philosophical writing. We do not think that laudable ends can shore up a muddled line of argument.

It is a problematic consequence of Stich's pragmatism that it does not appear to afford us the resources to identify sham reasoning. Perhaps Smith thinks that accepting some false belief will promote his ends of peace, love and harmony. Perhaps he is right. On Stich's view, given Smith's ends, Smith would have no reason at all to seek out evidence against his belief.<sup>25</sup> Furthermore, Smith would be justified in persisting in his false belief even in the face of disconfirming evidence. In such cases, however, I think the most natural response is to draw a distinction between different types of justification. Let us allow that Smith is *morally* justified in persisting in his false belief in the face of disconfirming evidence. At least, it would take serious argument to rule out the possibility. However, this would not mean that Smith is *epistemically* justified. For, intuitively, beliefs are epistemically justified to the extent that they are based on

---

<sup>22</sup> See Goldman (1999: 234–8) and especially Kunda (1990) for more on motivated reasoning.

<sup>23</sup> See Damasio (1997).

<sup>24</sup> I take the term 'sham' reasoning from C. S. Peirce. According to Peirce, sham reasoning occurs when 'it is no longer the reasoning which determines what the conclusion shall be, but it is the conclusion which determines what the reasoning shall be' (1931: 1. 57). He attacks sham reasoning as a corrupting influence on intellectual enquiry. See Haack (1998: 8–10) for further discussion.

<sup>25</sup> Smith might have a reason to seek out *prima facie* evidence against his belief simply in order to overcome it and fortify his original belief. Whilst this method of belief formation and maintenance is not unheard of in philosophy, it is of course just another form of sham reasoning.

epistemic reasons, which stand in favour of the *truth* of a proposition.<sup>26</sup> Peace, love and harmony may be reasons that bear on belief, but they do not affect the likelihood of the belief being true and they are not thereby epistemic reasons. The difficulty is that this natural distinction is not available to Stich because he has eliminated truth as the end of enquiry, and replaced it with the all-embracing category of everything an agent values. For Stich, belief in a proposition is justified *tout court* in light of an agent's values. As there is no longer any distinctively epistemic criterion of justification, it becomes impossible to say that a belief is epistemically justified, but not morally justified, or *vice versa*.<sup>27</sup>

For these reasons, it seems to me that Stich's account is unsatisfactory. Nevertheless, establishing a connection between the end of epistemic activity and what an epistemic agent values or desires is a common strategy for explaining epistemic normativity within a naturalistic framework. Whilst Stich is unusual in denying that the property of being true is relevant to the evaluation of beliefs, naturalists frequently maintain that the value of true beliefs is contingent upon non-epistemic ends. For instance, Kornblith offers an account of epistemic norms as 'universal hypothetical imperatives' (2002: 157). On Kornblith's view, as on Stich's, it is important to have beliefs that enable us to satisfy our goals. However, Kornblith emphasises the importance of *true* beliefs as a means to that end. Although I shall later question whether instrumentalists are entitled to appeal to true beliefs in this way, Kornblith's view does seem at first blush to have the merit of retaining the distinctively epistemic aspect of epistemic justification.

For Kornblith the value of true beliefs derives from their contribution to decision-making. He envisages choosing between available courses of action as involving a process that is appropriately modelled by cost-benefit analysis. Thus, he describes the homely example of choosing to buy one toaster or another:

---

<sup>26</sup> I defend this view with appropriate detail in chapter seven.

<sup>27</sup> See, however, Horwich (2006: 351) who takes the position that whilst truth has non-instrumental value, the value is a species of moral value.

We must figure out the consequences of the two purchases; we must assign values to each of them; we must do some arithmetic. (2002: 155)

So, when we decide how to act we multiply the expected value of a set of outcomes by the probability of the outcomes obtaining. In order to succeed in this piece of prudential arithmetic, Kornblith argues that we will typically need to have true beliefs:

It is thus of the first importance that our cognitive systems remain suitable for the purpose of performing the relevant cost-benefit calculations. And what this requires is that our cognitive systems be accurate, that is, that they reliably get at truth. (2002: 158)

Thus, Kornblith looks to yoke naturalised epistemology to the idea of subjective expected utility within decision theory.<sup>28</sup> Indeed, this will strike many as an attractive coalition of theories (and strike many others as an unholy alliance). For instance, it is one of the supposed attractions of decision theory that questions concerning final ends, traditionally the subject of philosophical enquiry, are replaced by more psychologically tractable questions concerning the ends we in fact have.<sup>29</sup>

Another apparent advantage of Kornblith's approach concerns motivation. As the ends of instrumental reasoning are by definition valuable to us, the mystery concerning the normative force of epistemic norms is supposed to be removed. At least we have a plausible looking explanation of motivational force. As Kornblith writes:

Someone who cares about acting in a way that furthers the things he cares about, and that includes all of us, has pragmatic reasons to favor a cognitive system that is effective in generating truths, whether he cares about the truth or not. (2002: 156)

Because we value the end, we will also value the means to that end; and here the means are true beliefs and reliable methods for acquiring them. As true beliefs

---

<sup>28</sup> See Hollis and Sugden (1993) for insightful criticism of decision theory. I shall briefly return to the subject in chapter eight.

<sup>29</sup> I say 'supposed' attraction because I argue against this sort of ant-rationalism in section 8.2 below.

are a precondition for the sort of cost-benefit analysis necessary for attaining whatever we value, we will always value truth.<sup>30</sup>

David Papineau (2003a) offers a very similar account to Kornblith's. Once more, for Papineau the aim of epistemic activity is truth, and patterns of reasoning which are reliably conducive to truth have normative force because true beliefs enable us to satisfy our desires. Thus, he writes:

On my view, what makes it the case that you ought to judge in certain ways on specific occasions is that this will be a means to your judging *truly*. There is nothing circular about this analysis, provided truth itself can be analysed without appealing to norms of judgement, and the adoption of truth as an aim is in turn explained by reference to moral or personal value attached to truth. (2003a: 20)

Moreover:

There is always a species of derived personal value to having true beliefs that are relevant to action, for such truth will always help you to find a way to satisfy whatever desires you have. (2003a: 17)

So, as with Kornblith's account, it is supposedly no mystery that the normative force of epistemic norms is felt universally. For, it is supposedly no mystery that we are motivated by our desires.<sup>31</sup>

Interestingly, Papineau also proposes that the value of true beliefs derived from non-epistemic goals allows naturalists to avoid certain difficulties raised by the rule-following considerations (2003a: 19–20). We might recall from chapter one that whilst the rule-following considerations primarily bear upon determining the truth conditions for propositions about meanings, much of the difficulty lies in identifying truth-makers with normative force – or, what I called value-makers. Consequently, Papineau is right to identify 'the question of why we ought to judge so-and-so on specific occasions' (2003a: 19) as an important issue stemming from Kripke's argument. Papineau's answer, in line with his account

---

<sup>30</sup> Although truth is 'pre-eminent', other epistemic values such as simplicity and conservativeness still have their familiar roles to play in epistemic justification insofar as they too are necessary for cost-benefit analysis. As Kornblith notes, 'a system of evaluation that was perfectly accurate but could not perform its evaluations in real time would be of little value' (2002: 158–9).

<sup>31</sup> I challenge this view in chapter eight where I discuss and reject desire-based theories of reasons.

of epistemic normativity, is that we should make arithmetical judgements in accordance with plus, rather than in accordance with a deviant quus-like function, because doing so pays off in terms of our further aims.

Although he does not fill in the details, it is not hard to see why Papineau might tend towards this sort of solution. It is, of course, quite natural to think that we ought to judge, and assert, that  $68 + 57 = 125$  because to do so is to judge truly. However, if we are to believe Kripke's interpretation, Wittgenstein's arguments show that we must replace truth conditions for sentences with assertability conditions, or, as Kripke also refers to them, 'justification conditions' (1982: 74). Kripke goes on to explain the justification conditions of an utterance in terms of its utility within a social practice (1982: 75–7).<sup>32</sup> This is supposed to be the 'sceptical solution' to the 'Wittgensteinian paradox'. I do not in fact think that this is an adequate solution, not least because most of this chapter will be devoted to arguing against instrumentalist accounts of epistemic normativity. Nevertheless, Papineau's contention suggests that the discussion is proceeding in the right direction and that I am not simply talking past naturalist concerns.

### *5.3 Categorical Epistemic Norms and Trivial Knowledge*

Advocates of naturalised epistemology frequently commit themselves, then, to a teleological theory of epistemic normativity. Truth-conducive norms of reason are valuable and motivating insofar as they promote the ends we have. However, a straightforward analogy with ethical theory suggests a possible concern resulting from the different normative properties of rules and ends. Ethical deontologists, for example, maintain that an action is justified insofar as it conforms to some categorical norm, and are then left with the tasks of explaining how we come to have knowledge of such apparently queer entities and explaining how they bind us. By contrast, advocates of teleological ethics, such as consequentialism or eudaimonism, argue that an action is good insofar as it is, or can reasonably be expected to be, conducive to some end, such as the greatest

---

<sup>32</sup> See, however, Winch (1983: 402–3) who takes Kripke to task for failing to distinguish between two ways in which it might be necessary to consider usage in order to identify the meaning of an expression. On the one hand, we might need to look at how an expression is used. On the other hand, we might need to look at the utility of an expression, or what use it is. Winch argues that Kripke wrongly attributes the latter position to Wittgenstein.

happiness of the greatest number or a flourishing life. The instrumentalist account of epistemic normativity appeals to a *telos* of epistemic practice as a source of normativity and so falls into the second category. However, one worry facing teleological accounts of normativity in general is the *prima facie* existence of categorical norms. For example, it is sometimes objected that because utilitarians hold that an action is only ever contingently right, they must be willing to countenance acts that strike us as unconditionally obligatory or impermissible. For instance, it appears morally impermissible to take an innocent man from the street for the purpose of forced organ donation. This intuition tends to hold despite the apparent utility of the action.<sup>33</sup> The thought is that the injunction not to take innocent life is categorically binding, regardless of any instrumental value attached to disregarding it.

Although this is only a bare sketch of a controversy, it suggests that there is a concern to be addressed. Moreover, if this is how matters stand in the moral case, there may be more pressure on those who deny the categorical force of epistemic norms. Anticipating the next two chapters, I should say that my sympathies are with the view that certain epistemic norms, i.e. those of basic logic, are categorically binding and that truth is a final end. In my view, rather than it being the case that our reasoning ought to conform to logic in order that we are better able to achieve our desired ends, part of what it is to be an epistemic agent is to be subject to certain logical norms and for one's belief formation to be regulated by the end of truth. One cannot be an epistemic agent unless one is rational in the sense of being accountable to and responsive to epistemic reasons. A non-rational creature or device, by contrast, can only passively record representations of the world. Consequently, it cannot be justified or unjustified in doing so. Thus, I agree with Emer O'Hagan when he writes:

It is not possible for an agent to engage in the practice of reasoning without being accountable to rational norms since such accountability is constitutive of agency. (2005: 42)

---

<sup>33</sup> Utilitarians do, of course, have resources to draw upon here. For example, they might point out that the disutility to a society in which people lived in constant fear of being snatched from the street is likely to outweigh the utility of saving the lives of a relatively small number of people needing organ transplants.

I also agree with Thomas Nagel:

In order to have the authority it claims, reason must be a form or category of thought from which there is no appeal beyond itself—whose validity is unconditional because it is necessarily employed in every purported challenge to itself. (1997: 7)

If this is right, norms of judgement are internal to epistemic practice, not external, as instrumentalists would have it. Hence, a rational agent would still be bound by epistemic norms even if violating them promoted his non-epistemic ends. Although non-epistemic reasons can trump epistemic reasons, they are not the source of epistemic normativity and, so, cannot negate them. As I shall defend these claims in the following chapters, the concern that instrumentalists fail to account for the categorical force of certain epistemic norms will strengthen in retrospect. Nevertheless, the above sketch at least suggests that it is coherent and plausible to think that there are categorical norms of reason.

Papineau responds to this sort of concern by asking us to consider a case in which it intuitively seems that we ought not to conform to the norms of judgement. The case is simply described and suggests that there are no categorical, *sui generis* norms bearing on belief formation. We are invited to consider the case of an elderly man who, aware that knowledge of a real probability of developing cancer would cause him considerable upset, arranges matters so as to avoid any such distressing evidence (2003a: 14). His behaviour is not obviously blameworthy and Papineau further claims that he does not seem to be ‘violating any prescriptions at all’ by adopting this ‘sensible’ strategy (2003a: 16). He describes the elderly man’s behaviour as ‘quite proper’ (2003a: 15).

Thus, Papineau sets up a thought experiment and gambles on our intuitions going in a particular direction. Although this is a perfectly reasonable strategy, I think that the particular example is unhelpful.<sup>34</sup> Because of the significance of the

---

<sup>34</sup> The reliability of philosophical intuitions has attracted considerable attention recently, largely as a result of so-called ‘experimental philosophy’ or ‘X-Phi’ advocated by some naturalists (see, for instance, Knobe and Nichols (2008)). I do not, however, see that these studies should occasion general scepticism with respect to philosophical intuitions, as I do not see how folk

moral reasons bearing on the case, intuitions are likely to vary about whether there really are no epistemic norms in play, or whether they are over-ridden. Indeed, given that the case is sketchily specified, it is far from clear to me that elderly man wouldn't be well advised to find out the truth, however distressing. Perhaps it would be better to prepare psychologically and set his affairs in order (or to have his nagging doubts allayed) than to live in a fools' paradise. Perhaps he would even have an obligation to do so in certain circumstances.<sup>35</sup> I am unsure and, as a practical matter, I would be reluctant to offer third party advice. Weighing such slight epistemic reasons against such important moral and prudential considerations is a formidably difficult task, rather like attempting to distinguish between fine wines adulterated with vinegar. Small differences in epistemic value are likely to be swamped.

So, I have reservations concerning the likely reliability of our intuitions bearing on the particular case. I do not, however, intend to challenge the claim that it is sometimes in our overall best interests to have false beliefs.<sup>36</sup> Instead, I would like to propose an alternative way to understand the case. Even if we allow that there is no epistemic norm telling against the elderly man's strategy, it is not clear that this establishes Papineau's point. For it is not clear that Papineau is describing a case of epistemic behaviour at all. Rather it might seem better to describe the man as *refraining* from epistemic activity with regard to the subject of his developing cancer. We do, after all, decline to investigate the truth of a great number of propositions without this implying that we thereby violate any norms of judgement. We might do so because the propositions are trivial, or because the investigation would be inappropriately time-consuming, or for

---

judgements undermine the intuitive judgements of careful, reflective philosophers. We do not, after all, appeal to the ballot box in order to settle philosophical disputes. I am also persuaded by Williamson (2005) that, at least very often, there is nothing especially mysterious about philosophical intuitions beyond an ability to handle counterfactual conditionals.

<sup>35</sup> See Lynch (2004: 131–6) for a general defence of the importance of self-knowledge to a flourishing life.

<sup>36</sup> Interesting in this context is the example of so-called 'depressive realism'. Taylor and Brown (1988) provide an overview of the (contested) phenomenon. In summary they write that 'considerable research evidence suggests that overly positive self-evaluations, exaggerated perceptions of control or mastery, and unrealistic optimism are characteristic of normal human thought. Moreover, these illusions appear to promote other criteria of mental health, including the ability to care about others, the ability to be happy or contented, and the ability to engage in productive and creative work' (1988: 193). See, however, Flanagan (2007: 170–81) for further discussion and criticism of the methodology of these studies.



moral, prudential or aesthetic reasons. For example, it would be unreasonable to dissolve a painting by Caravaggio in order to investigate its chemical composition. However, a decision to refrain from some activity does not constitute a rejection of the norms that govern it. If one decided that the aforementioned artistic destruction was a reasonable project, then one would be governed by the applicable epistemic norms, as in all one's epistemic endeavours.

Because examples requiring us to weigh different categories of reasons can be difficult to evaluate, I would like to consider an apparently simpler, less morally loaded, example, namely the case of unimportant or trivial knowledge.<sup>37</sup> The case is particularly interesting because trivial knowledge is frequently taken to count against the view that truth is of non-instrumental value. For it seems obvious to many philosophers that there are very many true propositions which we have no reason, let alone obligation, to come to know. This is taken, moreover, to demonstrate that there is no value to holding true beliefs except when they are instrumental to the satisfaction of our desires. For example, Stephen Grimm writes as follows:

If you propose an evening memorizing the phone book for Topeka, Kansas, and I decline, have I really missed an opportunity to enrich myself, from an epistemic point of view? If the truth is always intrinsically worth pursuing, then it seems that I have. And yet that conclusion seems ridiculous. (2008: 726)<sup>38</sup>

Contrary to these claims, however, I would like to argue that upon closer inspection cases of trivial knowledge provide very good evidence of the reverse, namely that the epistemic norm of truth applies regardless of instrumental considerations. Of course, partly for the reasons given above, this is not to say that we have an all-things-considered duty to pursue every truth. Nor is it even to say that we have an all-epistemic-things-considered duty to pursue every truth, for time is limited and many avenues of epistemic endeavour appear

---

<sup>37</sup> Similar considerations may apply to cases of knowledge which cannot even in principle further our non-epistemic interests. For example, knowledge of non-causal entities or entities that are causally isolated from our world.

<sup>38</sup> Zagzebski (2003: 20–1) and Pritchard (2007: 102) also appear to endorse the view that trivial knowledge is without value.

unpromising. We can reasonably judge in advance that many true beliefs are unlikely to be epistemically fruitful in the sense that they are unlikely either to enable the acquisition of further true beliefs or to promote understanding by laying bare the connections between beliefs. Instance of trivial knowledge are very likely to be epistemic dead ends.<sup>39</sup>

With these qualifications in mind, let us consider a particular example, say, the number of blades of grass on Parliament Square in London. Let us also agree for the sake of argument that this is a question about which we feel entirely indifferent. We do not value the answer, and, furthermore, there is no prospect that it will help us to attain any of the goods that we do in fact value. It is trivial knowledge without instrumental value. I take it, nonetheless, that the fact that we do not care about the matter does not simply settle the matter of whether we have reason to care. Imagine, for instance, that despite our complete indifference, we encounter reliable and conclusive evidence that there are precisely 100,000,000 blades of grass on Parliament Square. A meticulous team of philosophically inclined protesters assures of the fact (they have also resolved any questions of vagueness). Under such circumstances, I suggest that it would be irrational to fail to form a true belief about this matter of fact and continue to profess ignorance or to falsely believe that the answer was some other figure. For, according to one plausible view, rationality is, *inter alia*, a matter of recognising, and responding appropriately to reasons.<sup>40</sup> And, in this case, the testimony constitutes a decisive reason for belief.

Moreover, we should notice that the usual norms of reason apply to trivial beliefs. For example, we know that if there are 100,000,000 blades of grass on Parliament Square there are not 200,000,000 blades of grass. We know that the answer to the question of how many blades of grass are on Parliament Square is a number and not, say, a colour or an adverb. We know that the number of blades

---

<sup>39</sup> Indeed, it is difficult to rule out the possibility that there are some true propositions which should *never* be known all-things-considered. I have in mind examples such as propositions concerning destructive technologies or fundamentally demoralising facts concerning the human condition.

<sup>40</sup> Rationality is sometimes used in a highly restrictive sense to mean the rationality of finding and adopting means to some independently given end. I discuss the limitations of this sort of instrumental rationality in chapter eight.

of grass is either even or odd. We also know that the number is a good deal higher than one, and a good deal lower than the number of blades of grass in the whole of London. If the instrumentalist view were correct, however, there would be no epistemic norms bearing on our judgement in these matters, for each of these beliefs would be irrelevant to the satisfaction of our desires.<sup>41</sup> The fact that some beliefs are true and based on good reasons would not even weigh in the balance. As Papineau writes, the ‘prescription [to believe the truth] only applies to those beliefs that are relevant to your actions’ (2003a: 17). In fact, however, epistemic norms apply to beliefs quite generally and not only to instrumentally valuable beliefs. We cannot simply excuse ourselves from the demands of reason when they are irrelevant to our purposes.

The example of trivial beliefs suggests that we are subject to *sui generis* epistemic norms regardless of our desires. Although Papineau finds this counterintuitive, it seems to me that the view coheres well with our background intuitions. John Dewey once wrote that ‘man naturally prizes knowledge only for the sake of its bearing upon success and failure in attaining goods and avoiding evils’ (1929: 51). I think Dewey was wrong. Not only are we subject to epistemic norms whether or not we care about them, we are also naturally curious beings who tend to prize knowledge for its own sake. One indication of this is the thought that if of one had unlimited time and mental capacity, it would be a good to possess as much trivial knowledge as possible. Knowing every true proposition about the world except the number of blades of grass in Parliament Square is intuitively a worse state than omniscience. As ignorance is an epistemic imperfection, the possession of trivial knowledge always has some value.<sup>42</sup>

---

<sup>41</sup> It might be objected here that these are not trivial matters, but instances of very general considerations which have pragmatic significance elsewhere in our lives. This is closely related to the kind of rule-consequentialist response, which I come to in the next section.

<sup>42</sup> The method I am employing here to establish the non-instrumental value of true beliefs is very similar to G. E. Moore’s ‘method of isolation’ (*PE* 187–8). Perhaps Moore’s clearest use of the method is his attempt to establish that the value of beauty is not merely instrumental to the pleasure to be obtained from its appreciation (*PE* 83–6). To this end, Moore invites us to consider two worlds, one of superlative beauty, the other of supreme ugliness. Both worlds are permanently devoid of humans or other beings capable of aesthetic judgement. Moore argues that if we are inclined to view the former world as the more valuable, it must be because beauty has non-instrumental value.

Michael Lynch also provides an example which appears to show that we value true beliefs, irrespective of their instrumental value:

Suppose that, unbeknownst to us, the world began yesterday – it *seems* older, but it isn't. If I really lived in a Russell World, as I'll call it, almost all my beliefs about the past would be false. Yet my desires would be equally satisfied in both worlds. This is because the future of the two worlds unfolds in exactly the same way ... Yet given the choice between living in the actual world and living in a Russell World, I strongly prefer the actual world. (2004: 18)

My intuition is the same as Lynch's. Although these beliefs about the past are not trivial, they share with trivial knowledge the feature of not connecting up with any of our non-epistemic desires. Nevertheless, I would strongly prefer not to live in a Russell World.<sup>43</sup> Furthermore, I suspect that very few people would be content to decide the matter on the flip of a coin, as one should expect if it were a matter of indifference. We do in fact value true beliefs, or getting things right, independently of utility. Examples such as Papineau's in which the value of truth is swamped by other considerations do not tend to show otherwise. Moreover, contrary to a common view of the matter, examples of trivial knowledge tend to reinforce the thought that truth is valuable for its own sake.

#### *5.4 Instrumentalism, Contingency and Truth*

According to instrumentalists, epistemic norms are binding insofar as conforming to them makes it more likely that we will form beliefs that promote our ends. Papineau and Kornblith both argue that beliefs are more likely to promote our ends if they represent the world as it is, thus making true beliefs contingently valuable proximate goals. In response, I have motivated the thought that certain norms of reason are categorical and I have also argued that true beliefs have non-instrumental epistemic value. To this end, I have argued that examples such as Papineau's elderly man and cases of trivial knowledge do not tell against true beliefs possessing non-instrumental value. Indeed, I have suggested that cases of trivial knowledge indicate the reverse. In this section, I would like to present several further worries for the view that epistemic norms can be reduced to instrumental norms. In particular, I would like to argue that it

---

<sup>43</sup> Or be a brain in a vat, or the victim of a Cartesian Demon or subject to any other form of systematic illusion.

is far more difficult than instrumentalists suppose to find an instrumental role for truth when the final goal of belief formation is non-epistemic.

Instrumental reasons are not only contingent upon having a particular end, but also upon having further beliefs. In particular, they are only rationally compelling if accompanied by the belief that the course of action is actually a means to the end in question. For instance, the prescription ‘if you value car-speed, then you ought to tune your car well’ is only a reason to act if you believe that tuning your car well will make it go faster. By contrast, the prescription ‘if you value car-speed, then you ought to tune your violin well’ has no rational force because it is apparently predicated upon a false belief. We have no reason to believe that there is any connection between engine performance and well-tuned violins. So, given that instrumentalists recommend that we form true beliefs as a means to pragmatic ends, it needs to be the case that a reliable means-end relationship holds between the two.

We have already touched on one reason for doubting that the possession of true beliefs tends to promote pragmatic ends. This was Stich’s argument that true beliefs are no more likely to get us what we want than beliefs instantiating any number of nearby truth-like functions.<sup>44</sup> Although I have objected to Stich’s account of epistemic normativity, his concerns about the relation between true belief and successful action merit independent consideration. It could be, for instance, that true beliefs have final epistemic value, but, for just the reasons Stich offers, no instrumental value with respect to the satisfaction of non-epistemic goals. This would constitute a decisive reason to reject the type of truth-tropic instrumentalism defended by Papineau and Kornblith.

Although Stich’s arguments constitute an interesting area for exploration, I do not propose to study them here in any detail.<sup>45</sup> I would, however, like to develop

---

<sup>44</sup> Stich takes this to show that we have no particular reason to value true beliefs. A more appropriate moral, it seems to me, would be that we have no reason to value true beliefs in particular. If Stich’s argument were sound, true beliefs *and* beliefs instantiating nearby truth-like functions would have instrumental value.

<sup>45</sup> One reason for my reluctance is that Stich’s arguments depend on a number of substantive claims about the metaphysics of belief, which I am not prepared to defend. See Goldman (1991) for further criticism along these lines.

one strand of Stich's discussion. In particular, I would like to pay attention to the nexus between mental states on the one hand, and action on the other. Consider, in this light, the following statement from Patricia Churchland:

Boiled down to essentials, a nervous system enables the organism to succeed in the four F's: feeding, fleeing, fighting and reproduction. The principal chore of nervous systems is to get the body parts where they should be in order that the organism may survive. Insofar as representations serve that function, representations are a good thing ... Truth, whatever that is, definitely takes the hindmost. (1987: 548–9)

Churchland is pointing out that evolutionary pressures bear directly on behaviour and only indirectly on representations. From an evolutionary perspective, Churchland suggests, representations should be evaluated against the standard of fitness rather than against some poorly understood notion of truth. Furthermore, even though we need representations of our environment in order to act, there are many ways in which a representation could be fitness enhancing other than through its truth or accuracy.<sup>46</sup> In the evolutionary case, for instance, it is very likely that we have inherited a number of quick and nasty cognitive strategies, which are adaptive despite frequently producing false beliefs.<sup>47</sup> Indeed, this is one plausible explanation for many of the psychological findings concerning systematic biases in human cognition.<sup>48</sup> Time is short, cognitive resources are limited and it is sometimes better to rely on rules of thumb than careful reasoning.

Although Churchland is concerned with evolutionary fitness, her observations apply to goal-orientated behaviour quite generally. Successful action is neither determined by nor always depends upon true beliefs. Rather, it depends on the interplay between a number of mental, behavioural and environmental factors. Consequently, there are always ways in which an action predicated on a false belief can succeed. Indeed, we can vary the factors around the nexus of thought

---

<sup>46</sup> Rorty (1995: 282–3) also thinks that evolutionary considerations scotch the idea that there is any particular connection between truth and pragmatic value.

<sup>47</sup> For instance, Dennett (2006) offers an explanation of the origins of religious belief along these lines. He points out that whereas failing to detect a predator (or a potential mate) can be a fatal error, a false positive is likely to cost very little. Consequently, it is adaptive to err on the side of caution and identify living creatures on the basis of slender evidence. This is an instance of what is sometimes called the 'Garcia effect'. Dennett suggests that it might explain primitive forms of animism upon, which other religions have built.

<sup>48</sup> See, for instance, Nisbett and Ross (1980), Stich (1985), Wilson (2002) and Papineau (2003b).

and behaviour in order to construct any number of examples of successful action without true belief. In this vein, Plantinga (1993b: 226; 2002) highlights examples of gerrymandered belief-desire systems which produce adaptive behaviour. For instance, a desire to be eaten alive could be combined with the belief that the best way to approach a tiger is to run away from it.<sup>49</sup> Of course, it is difficult to see how these gerrymandered belief-desire systems could be adaptive in the long run.<sup>50</sup> I shall come to general policies shortly. But my interest here is not to argue against the view that evolutionary processes tends to produce true beliefs, any more than I want to argue against the view that true beliefs are more likely than false beliefs to lead to the satisfaction of our desires. Rather, I am attempting to bring out the structure of the instrumentalist position in order to focus more clearly on the role that truth plays.

Churchland's proposal concerning the value of representations mirrors the instrumentalist's proposal that we evaluate processes of belief formation against the standard of pragmatic value. The key structural difference is that instrumentalists such as Kornblith and Papineau retain the proximate goal of truth. For Churchland, by contrast, truth is of little or no concern. If, however, the final goal of epistemic activity is not true belief, why don't instrumentalists follow Churchland in abandoning concern for the means? The answer is not that there is a necessary connection between true beliefs and pragmatic value. As we have already seen, true beliefs and pragmatic ends are contingently related at best.<sup>51</sup> Indeed, Papineau relies on this fact when he argues that there are no categorical norms of judgment applying to the elderly man in fear of the possibility of cancer. I have of course argued that Papineau is mistaken in believing that there are no epistemic norms bearing on cases in which true beliefs do not tend to satisfy our desires. Nevertheless, the question remains, if epistemic normativity is not finally grounded in truth, what is the benefit of relating epistemic norms to the proximate aim of truth?

---

<sup>49</sup> This is in the course of Plantinga's 'Evolutionary Argument Against Naturalism'.

<sup>50</sup> Ramsey (2002: 20–5) argues this point at length.

<sup>51</sup> Maddy (2007: 155–6) describes a particularly good example of this contingency, namely Joseph Priestley's false beliefs concerning phlogiston successfully leading to a method for the production of oxygenated air.

One possible answer would involve shifting the focus from individual beliefs to rules or general epistemic policies. Even if we would be better off having false beliefs in particular cases, perhaps a general policy of forming true beliefs would satisfy our desires better than a case-by-case policy. This looks quite plausible and, if correct, it would seem to provide a clear role for true beliefs within an instrumentalist account. The problem is that it is difficult to see why we should evaluate general policies in the light of promoting our desires, but not evaluate individual beliefs in the same way. It is analogous to the difficulties indirect- or rule-consequentialists face in motivating their position. If the ultimate source of normativity is pragmatic value, then there is no obvious reason why an individual belief should not be evaluated according to that standard. It appears *ad hoc* to invoke rules or general policies when it is ultimately the consequences of holding true beliefs that matter. If, on the other hand, it is the epistemic rules that matter, we have a completely different type of theory, namely a deontological account.<sup>52</sup>

It might be suggested that even if any particular true belief is not necessary for fulfilling our ends, and even if it is not possible to appeal to general policies of belief formation, it is still more valuable when our ends are in fact fulfilled on the basis of true beliefs. After all, it seems better on the whole to proceed on the basis of knowledge than to rely on chance or gerrymandered belief-desire systems. Although I think this is right in a certain sense, I do not think that the observation helps instrumentalists to establish a role for truth in their account of epistemic normativity. For, as a general principle, it is not true that a means provides a reason for action in addition to the end it serves. If, for example, I need a library book, I do not have two reasons to visit the library, namely my need for the book and the fact that I must visit the library to obtain the book. Nor

---

<sup>52</sup> Hooker (2000) challenges the argument that rule-consequentialism collapses into act-consequentialism by denying that rule-consequentialists have an overarching commitment to maximising the good. Rather, he suggests that the commitment is to do what is impartially defensible (2000: 101). This is interestingly similar to Parfit's recent defence of rule-consequentialism on Kantian grounds. Thus, Parfit writes: 'Everyone ought to follow the principles whose being universal laws would make things go best, because these are the only principles whose being universal laws everyone could rationally will' (2011: 418). Although the further details of the debate are complicated, I do not see how instrumentalists could construct a parallel defence. For, once the overarching commitment to acting in a way that satisfies desires is abandoned, I do not see how naturalists could motivate a commitment to believe what is impartially defensible or, indeed, how they could explain what it is to be impartially defensible. Perhaps, though, this is just a lack of imagination on my part. For further discussion of the ethical case, see Hooker's (2007) exchange with Card (2007).



is obtaining the book more valuable for the fact that I have successfully visited the library in order to acquire it. Insofar as the value of obtaining the book is concerned, it would be just as well if it fell into my lap from the heavens.<sup>53</sup>

Furthermore, it does not matter here if the means is reliably related to the ends. Consider the following example from Linda Zagzebski:

A reliable espresso maker is good because espresso is good ... The good of the product makes the reliability of the sources that produces it good, but the reliability of the source does not give the product an additional boost of value. The liquid in the cup is not improved by the fact that it comes from a reliable espresso maker. If the espresso tastes good, it makes no difference if it comes from an unreliable espresso machine. (2003: 13)

Although Zagzebski is highlighting a particular difficulty for reliabilism, her point depends on a general principle concerning the transmission of value between means and ends. The example is intended to show that the fact that a true belief has been produced by a reliable cognitive process does not make it any more valuable than if the true belief had been produced in an unreliable way.<sup>54</sup> As with my example of the library book, all the value comes from the end which is the source of normativity. Although the means is instrumentally valuable insofar as it produces the end, it is not an additional source of value. When the value of a good is transmitted from means to end, we do not benefit twice over from the means and the end. To think otherwise would be double counting.

We can apply this point to instrumentalism as follows. Instrumentalists claim that the proximate goal of truth is valuable because it reliably enables us to satisfy our non-epistemic ends.<sup>55</sup> Very likely this is so. For example, the prospect of developing a vaccine is an excellent reason to form true beliefs concerning the genetic code of a particular strain of the influenza virus.

---

<sup>53</sup> There are of course other reasons why the particular means can affect the value of the means-end whole. For instance, the means could be wicked, or involve the exercise of virtue or not worth the candle. But these do not result from the means-end relationship itself.

<sup>54</sup> This is a problem for reliabilism because it challenges the idea that the justification condition for knowledge can be met by the reliability of the process which results in a true belief.

<sup>55</sup> Of course, we might (and do) also desire to have true beliefs. Papineau (2003b) emphasises this fact and tells an evolutionary story to explain our curiosity. However, as long as the normativity is grounded in the desire for truth rather than the truth itself, this is no more than a variation on a theme.

However, if the instrumentalist view were correct, all of the value of the true beliefs would be explained by the value of the non-epistemic end. For, the lesson of the example above is that the means *per se* have no value or disvalue. We do not have the epistemic value of the means *as well as* the pragmatic value of the end. But, if no distinctively epistemic value results from the employment of particular means, it looks difficult to justify the instrumentalist's emphasis on true beliefs. It would be just as well if an agent's desires were satisfied through any other mechanism, such as Plantinga's gerrymandered belief-desire systems. It is hard to see, then, how the instrumentalist's model is supposed to explain *epistemic* value, when the value is entirely accounted for by non-epistemic ends. It is not enough to claim that the value is epistemic simply in virtue of the fact that it is *beliefs* which are being evaluated. For, we can evaluate beliefs on moral, aesthetic and any number of other non-epistemic grounds.

A couple of paragraphs back, I did allow that it is better in a certain sense if we achieve our practical goals as a result of having true beliefs. This is because otherwise we are merely lucky, and, as any gambler will testify, luck rarely holds. Moreover, we do not deserve credit for satisfying our goals by chance. Any such success is not due to our agency and cannot be depended upon in future. By contrast, someone who satisfies her goals by skilfully acquiring true beliefs is likely to satisfy her future desires and those of others. Just as reliability is a desirable feature in an espresso maker, but does not affect the value of the finished product, so the knack of producing valuable outcomes is a desirable feature in a member of a community. But, although these facts are important when it comes to the appraisal of epistemic agents, they do not affect the value of true beliefs *per se*. If a true belief does not have epistemic value in virtue of enabling us to satisfy a particular desire, then neither do any number of true beliefs have epistemic value in virtue of enabling the wider community to satisfy any number of desires.

We have seen, then, that according to the instrumentalist picture particular true beliefs are not the ultimate source of epistemic normativity, and that they are uncertainly correlated with pragmatic value. We have also seen that an appeal to a more general epistemic policy of forming true beliefs is under-motivated and

appears to fall foul of a dilemma frequently posed of rule-consequentialism. Finally, we have seen that true beliefs do not add epistemic value beyond the non-epistemic, instrumental value they inherit from pragmatic ends. These conclusions call into question whether instrumentalists are entitled to appeal to true beliefs. And, if they are not entitled, it is unclear on what grounds they may be said to offer an account of epistemic normativity, as opposed to merely drawing our attention to the pragmatic norms which bear on belief.

As we saw in the previous chapter, the natural sciences are able to provide explanatory reasons for our cognitive practices. Such reasons are, however, logically distinct from normative reasons. They do not necessarily justify our cognitive practices. Instrumentalists, then, attempt to pass the buck by connecting our cognitive practices to a further end. The end of truth does provide the right sort of epistemic reason, which can be transmitted back to our cognitive practices thereby justifying them. However, the naturalist is unable to provide a satisfactory account of why the end of truth is normative, except in terms of our pragmatic goals. In this chapter I have tried to show that this provides the wrong sort of reason. In fact, I think the problem is much worse than this. As I shall argue in chapter eight, desires do not count as reasons. Consequently, instrumentalism offers us no real normativity at all.

Given these difficulties, it seems better to sever the link with pragmatic value and simply ground epistemic norms directly in truth as a final value. Indeed, this is the view I shall defend in chapter seven. So long as we are not trying to ground the normative force of true beliefs in further desires, the difficulties we encountered above do not arise. In particular, there is no tension between holding truth as the end against which epistemic activity is evaluated and recognising that truth is only contingently related to desire. In the following chapters I shall defend such an account. I should also like to hold onto the instrumentalist insight that there is an important category of normative force derived from the ends of a practice or mode of activity. Much of my work will consist in explaining how this can be a source of normativity without passing the buck or appealing to the inappropriate category of desire. I should also like to do justice to the intuition that there are categorical as well as hypothetical norms,

explain their origin and locate their value-makers. Finally, I should like to explain how these categories of norms fit together and why they escape a naturalistic worldview. This will be the burden of the following chapters.

## **Chapter Six**

### **A Sketch of the Good and the Right**

In previous chapters, I studied the tension between naturalism and normativity, the nature of naturalism and the limitations of certain naturalistic approaches to normativity. In this chapter I begin to construct a positive account of the nature of normativity. Some aspects of my view have already been stated or implied. I shall be defending an objective, externalist theory of reasons and I shall not be relying on a scientific perspective. I shall, however, retain the instrumentalist's insight that there is an important connection between acting well and achieving goals. But I would like to combine this with an explanation of the sense in which some norms are categorical. In later chapters, I shall explore questions surrounding normative force, motivation and normative epistemology.

I shall proceed by examining a number of important normative concepts. In particular, I would like to focus on the relationships between the good, the right and practices with their associated ends. In section 6.1, I look at the logical grammar of evaluative terms, paying particular attention to Geach's (1956) proposal that 'good' is an attributive adjective. In section 6.2, I look at the neo-Aristotelian idea that the relationship we can understand goodness in functional terms. In sections 6.3 and 6.4, I sketch an account of the good and the right respectively, using the game of chess as an example. Finally, I broaden out my discussion to consider the extent to which chess is analogous to epistemic and ethical activity. This leads to me to consider the nature of needs and finally the meaning or value of life as a whole. I approach these questions in section 6.5.

#### *6.1 The Logical Grammar of Evaluative Terms*

What is goodness? I would like to start with some comments about the logical grammar of 'good' and 'bad'. Grammarians distinguish between *attributive* and *predicative* adjectives. In his paper 'Good and Evil', Peter Geach (1956) drew philosophers' attention to this distinction and argued that it could be extended and applied to our understanding of the logic of evaluative terms such as 'good' and 'bad'. So, consider a complex predicate of the form 'is an AN' where A is an adjective and N is a noun. For example, 'is a red book', 'is a sharp spade' or

‘is a good move’. Now, we can draw a distinction between two types of adjectives by looking at the way complex predicates behave. In some instances the predication ‘x is an AN’ logically decomposes into the two predications ‘x is an N’ and ‘x is A’. So, using Geach’s example, ‘x is a red book’ logically decomposes into ‘x is a book’ and ‘x is red’.<sup>1</sup> In these cases, the truth-value of the complex predicate is simply the truth-value of the conjunction of its component predicates, i.e. ‘x is an N’ and ‘x is A’. If it is true that x is red and that x is a book, then it is true that x is a red book. When complex predicates behave in this way, we can say that the adjective is predicative.

However, not all adjectives behave like ‘red’. Some complex predicates are inferentially irregular in the sense that their truth-values are not given by the conjunction of their components. Importantly for present purposes, Geach observed that the predicates ‘is good’ and ‘is bad’ are inferentially irregular. Thus, the predication ‘x is a good book’ does not logically decompose into the predications ‘x is a book’ and ‘x is good’. Clearly, it fails in the case of the adjective ‘good’ rather than the noun ‘book’.<sup>2</sup> We can infer from the truth of the proposition ‘x is a good book’ that ‘x is a book’, but not that ‘x is good’. Or, consider the proposition ‘x is a good poison’. We clearly cannot make two unqualified inferences that ‘x is a poison’ and that ‘x is good’. Rather we mean something like ‘x is good as poison’ or ‘x has all the properties poison should have’ or ‘x is well-fitted to doing those things which poison is for’. In this sense ‘good’ and ‘bad’ are like ‘big’ and ‘small’. We cannot infer from ‘x is a big flea’ that ‘x is big’. Even big fleas are rather small creatures. Rather, we should understand something like ‘this is a big example of a flea’ bearing in mind that it is in the nature of fleas to be small by most standards.<sup>3</sup>

---

<sup>1</sup> Geach’s example is complicated by the fact that ascriptions of colour terms such as ‘red’ seem to be context dependent. For instance, ‘x is a white wine’ does not decompose into ‘x is white’ and ‘x is a wine’. For white wine is actually a yellowish liquid. For more on this point see Thomson (2008: 233–48).

<sup>2</sup> Rind and Tillinghast (2008: 85–6) argue that this is a contingent feature of natural languages and that it is possible to construct nouns that are inferentially irregular in the same way.

<sup>3</sup> Mightn’t we sometimes want to say that something is big without qualification, as, for example, when we contemplate a mountain and say “Goodness, that’s big!”? One response would be to insist that what we really mean is something like “Goodness, that’s big *for a mountain!*” But, of course, it might be rather small for a mountain and we might still be inclined to say “Goodness, that’s big!” It is natural, then, to think that what we mean is simply ‘Goodness, that mountain is a big *thing*’, employing ‘big’ as a predicative adjective. We are simply in awe of its size. This is

I think that Geach is right that the underlying logical form of ‘good’ and ‘bad’ suggest that they are attributive adjectives. Or, at least that this is the case when evaluative terms such as ‘good’ are used to qualify nouns. What this shows is that the predication ‘x is good’ is always, or very often, incomplete. Geach puts the point as follows:

Even when ‘good’ or ‘bad’ stands by itself as a predicate, and is thus grammatically predicative, some substantive has to be understood; there is no such thing as being just good or bad, there is only being a good or bad so-and-so. (1956: 34)<sup>4</sup>

Although I do not want to commit myself to Geach’s view that ‘good’ *always* functions in this way – I discuss a predicative use of the term in the final section of the chapter – it is clear that predicative uses of ‘good’ do not normally make sense as they stand. Philippa Foot (2001: 2) described how she sometimes secured recognition of this point when giving talks by holding up a small piece of torn paper and asking whether or not it was good. An offer to pass it round so the audience can get a better look provoked a laugh in recognition of the absurdity. It is meaningless to talk about scraps of paper being good in an unqualified sense. Similarly, Judith Jarvis Thomson offers the following example:

You are standing in front of the array of melons at your grocer’s, feeling helpless. Your grocer notices. He points to one in particular and says, “That one’s good.” ... It would be utterly astonishing if when you asked, “Do you mean that that’s a good melon?” he replied, “Oh dear me no, I haven’t the faintest idea whether it’s a good melon, I meant only that it’s a good thing. (2008: 13)

Again, this is comical and the comedy comes from the fact that it (generally) makes no sense to say that some thing is simply good.

---

plausible until we remember that mountains, even big ones, are not big things in comparison to, say, continents and that they are very tiny in comparison to planets, galaxies or the universe. What we really mean is something like ‘this is a big feature of the landscape’ or ‘this is big on the human scale of things’. Perhaps one could say of the universe that it is big without qualification, but it is not at all clear what that claim would amount to, for it is not clear what it would be for the universe to count as small.

<sup>4</sup> J. L. Austin employs the metaphor of ‘good’ being ‘substantive-hungry’. He also speaks of ‘good’ as ‘crying out for substantives’ (1962: 68–9). In this sense, Austin suggests, ‘good’ is similar to ‘real’. Something of the same idea can be found in Hare (1952: 133). See Hare (1957: 103 fn. 2) for a little more on the prehistory of Geach’s idea.

One way in which Geach explains the idea that ‘good’ is an attributive adjective is as follows:

I could ascertain that a distant object is a red car because I can see it is red and a keener-sighted but colour-blind friend can see it is a car; there is no such possibility of ascertaining that a thing is a good car by pooling independent information that it is good and that it is a car. This sort of example shows that ‘good’ like ‘bad’ is essentially an attributive adjective. (1956: 33–4)

Geach’s comments may provide a partial explanation of why evaluative properties are so difficult to accommodate in a naturalistic picture of the world. Geach is suggesting that the task of determining whether some thing is good cannot be parcelled out among friends. In this respect, evaluative properties seem like aesthetic properties. However, I emphasised in chapter three that scientific enterprise depends upon distributed cognition and is predicated upon the interchangeability of investigators. The problem is not that there are any spooky facts involved or that some special faculty of intuition is needed in order to discern whether a car is good. For, imagine that Geach has another friend who is an experienced mechanic and can hear that the car’s engine is running smoothly (and let us agree for the sake of simplicity that this is all there is to a car being good). In this case, I don’t see any objection to saying that it would be possible to pool the information that the distant object is a car and that its engine is running smoothly and conclude that it is a good car. It would be necessary, however, for knowledge of the empirical facts concerning the functioning of the engine to be combined in practical reason with an understanding of what makes a car a good one. Although there is nothing supernatural about these considerations, one cannot simply *see* what makes a car good in the same way that one sees that it is red.<sup>5</sup>

---

<sup>5</sup> Of course, the phenomenology of perception depends on one’s background knowledge and experience. An experienced mechanic may sincerely assert that she simply hears that a car engine is not running well. Similarly, strong chess players report ‘seeing’ that a move or position is ‘strong’ or ‘weak’. This helps to explain the relatively high standard of ‘blitz’ chess in which all the moves are made in less than five minutes and ‘positional sense’ is all-important. See further Chase and Simon (1973) for research into chess masters and perceptual processing. I return to these topics in chapter nine when I argue that emotional experience is analogous to perceptual experience in a number of ways.



The above considerations suggest an epistemological test for determining whether an adjective is predicative or attributive. However, it has occurred to a number of philosophers that there may be a further metaphysical explanation.<sup>6</sup> It is widely believed that goodness is a property.<sup>7</sup> Surface grammar suggests no difference between the predicate ‘is good’ and predicates such as ‘is red’ and ‘is spherical’. As it is normally the function of predicates to denote properties, it is natural to think that a sentence such as ‘x is good’ has the function of attributing the property of goodness to x. For instance, if properties are universals, then to say that ‘x is good’ is to say that x instantiates the universal of goodness. It is then very natural to think that the primary task of ethics is to identify the universal of goodness, or whatever property the predicate ‘is good’ denotes. Once we have done so, we can ask where this property is instantiated and how this is relevant to our conduct.

This view is certainly attractive and a great deal of philosophical writing has presupposed it, either explicitly or implicitly. Influentially, G. E. Moore held this sort of view of ethics. When he wrote that ethics is ‘the general enquiry into what is good’, (*PE* 2) he was using the ‘is’ of identity before the ‘is’ of predication. Thus Thomson (1997: 273) calls the sort of view I have been sketching ‘Moore’s story’, although it is very far from being peculiar to Moore. Thomson also suggests that Moore’s commitment to the view that goodness is a property had a baleful influence on twentieth century metaethics. For it is extraordinarily difficult to see what property could be exemplified by all of the things that are good or how we should go about identifying this property. In particular, locating the property of goodness in the natural order is a daunting task. Consequently, as Thomson puts it, the property of goodness is epistemologically and metaphysically ‘dark’ (2008: 11).

We have already approached these issues from one perspective during our discussion of the Open Question Argument in chapter one. I suggested that what makes normative properties especially difficult to identify with scientifically

---

<sup>6</sup> The following line of argument is strongly influenced by Thomson (1997) and (2008). See also Foot (2001).

<sup>7</sup> Or that it is a property *if it is anything*. Error theorists such as Mackie (1977) and Joyce (2006) argue that moral predicates fail to refer.

available properties is the phenomenon of normative force. Moore's own view, of course, was that goodness could only be a non-natural property revealed by a dedicated cognitive faculty of intuition. A more common response has been the non-cognitivist line that what is shared by all of those things we call 'good' is not some mysterious property of goodness, but the fact that we approve of them in some way. Thus, normativity is built into judgement rather than the world. If, however, Geach was right that 'good' and 'bad' are attributive adjectives it appears that Moore's story rests on a mistake. The fact that we are wont to say 'x is good' disguises the fact that 'is good' is an incomplete predicate. Much worse, it sends us off on a wild-goose chase looking for the property denoted by the predicate 'is good', such as the universal *goodness* that all good things instantiate. However, it is no more plausible to think that there is a property of goodness that all good things instantiate, than to think that there is a property of bigness instantiated by big fleas, big noses and big attics.<sup>8</sup> The metaphysical explanation, then, for the fact that we cannot pool our knowledge that 'x is a car' with our knowledge that 'x is good' to get the conclusion that 'x is a good car' is that there is no property of goodness for the predicate 'is good' to denote and, so, no truth-maker for the proposition 'x is good'. If this is right, it is little wonder that Moore struggled to find a real definition of the good.

Saying that there is no property of goodness is not to say that the term 'good' is meaningless or that there is no property denoted by particular attributive uses of the term. Nor is to say, with non-cognitivists, that there is nothing in virtue of which propositions including the term 'good' are true. Nor, finally, is it to concede that the term 'good' involves us in a hopeless hotchpotch of equivocation. We are still free to consider how attributive uses of 'good' function and under what conditions it is true to say that something is good. I approach this task below, but with the qualification that we should not assume that there is single, legitimate way in which 'good' functions in English, let alone a single way in which cognate and near cognates of 'good' function in other

---

<sup>8</sup> Cp. EN 1096b23–26.

languages.<sup>9</sup> Nevertheless, it would be an important step to identify a central and widespread use of the term with normative implications.

One way in which we might understand how the term ‘good’ functions is suggested by J. O. Urmson’s (1950) observation that ‘good’ is used primarily as a grading or evaluative label. To say that something is good is normally to say that it measures up more or less well to some standard or set of criteria. To say that an apple is good is to say, *inter alia*, that it has a pleasing taste, that it is pest-free and relatively blemish-free and that it stores well. To the extent that an apple meets these criteria, it is appropriate to say that it is a good apple. To the extent it falls short of these standards, it is a bad apple. Although not essential to grading, approximation to a standard also allows us to make comparative judgments.<sup>10</sup> Thus, Urmson pays great attention to the classificatory scheme of ‘super’, ‘extra fancy’ and ‘domestic’ grade apples, detailing their particular characteristics (1950: 151–4). In this way evaluative judgments contrast with deontological judgments. We might say that one thing is better, more excellent or closer to the ideal than another, but not ‘righter’ or more obligatory.

The etymology of ‘normativity’, from the Latin ‘*norma*’ meaning ‘builder’s square’, also suggests a close connection between the idea of a norm and the idea of measuring up to a standard. How, though, do we determine the appropriate standard against which to measure? Well, a starting point is that it depends on the kind of thing we are evaluating. This was part of Geach’s point in saying that a substantive must be understood when we make attributions of goodness. The proposition ‘x is good’ is typically elliptical for the proposition ‘x is a good K’

---

<sup>9</sup> For more on this point, see Wiggins (2009: 195–6).

<sup>10</sup> J. L. Evans objects to Urmson’s account on the grounds that it cannot deal with cases in which there is only a single specimen of a kind. In such cases, Evans writes: ‘we could not compare it with anything and, therefore, we could not say it was good in the sense of being good of its kind. There can be no superlative judgments unless there can also be comparative judgments’ (1962: 30). This shows a misunderstanding. When we grade an example of a kind the comparison is not with other members of the kind, but with an independent standard. Although we grade apples (and students) in order to rank them, grading is not essentially concerned with intra-class comparisons. It could be that *all* members of a kind are equally excellent, or far from the ideal. What may explain Evans’ objection is the thought that we would struggle to evaluate a single instance of a kind we had not previously encountered. But this is no different from our difficulty in evaluating an antique artefact when we do not know what it was for. It would not help, however, to encounter a vast trove of antique artefacts with minor differences. We do not need a comparison class, but an understanding of what constitutes the relevant standard of evaluation.

where K is the kind to which x belongs. What makes a good apple is quite different to what makes a good car. For example, we shouldn't expect an apple to withstand the impact of human incisors. These observations suggest a two-step procedure for evaluating propositions of the form 'x is a good K'. First, we identify the relevant kind K. Second, we measure x against the normative standard governing things of kind K. Sometimes, however, the two steps are conceptually inseparable. Identifying an object as belonging to a particular kind can depend on grasping the normative criteria that govern membership of the kind. For instance, part of what it is to identify an object as a spade is to identify it as a digging implement.<sup>11</sup> 'Spade' is a success term in the sense that an object that is not *at all* suitable for digging could not count as one.<sup>12</sup> Although there are such things as rusty, blunt and lightweight spades, there is no such thing as a spade made out of room temperature blancmange.

It might be objected that this still does not tell us how to determine the particular criteria against which to evaluate a specimen of a kind. Furthermore, many kinds have no plausible normative standards against which they can be evaluated. For example, it seems that there is nothing that it is to be good mud (although there is such a thing as good mud for the purposes of farmers, brick-makers and bathing hippopotamuses). Put another way, there is no property of being good *qua* mud, although mud may have the property of being good *qua* something else, such as building material. So, some kinds, such as spades, seem to come with normative standards built into their identity conditions.<sup>13</sup> Other kinds, such as mud, do not. Ideally, we would want a principled way to distinguish the one from the other.

## 6.2 Functions and Natural Goodness

---

<sup>11</sup> The thought here can be traced back to Aristotle's *ergon* argument (*EN* I: 7). The Greek term '*ergon*' is ambiguous in a potentially suggestive and helpful manner, meaning both 'function' and 'characteristic work'. So, in order to determine a thing's *ergon*, we must first consider what sort of work or activity is characteristic of it. This means understanding its form (*eidōs*). As Aristotle writes in the *Meteorologica*: 'What a thing is is always determined by its function: a thing really is itself when it can perform its function; an eye, for instance when it can see' (390a10-2). Hence, the ideas of a thing's function and a thing's identity come together in the idea of a thing's *ergon*.

<sup>12</sup> Cp. *De Partibus Animalium* 640b36-641a3 and *De Anima* 412b21-3; 416a3-6.

<sup>13</sup> As Thomson puts it, some kinds are 'goodness-fixing' (2008: 21).

One possibility is to draw on teleological considerations at this stage. As a number of philosophers in the Aristotelian tradition have observed, there is a close relationship between the goodness of a thing and its purpose or function. A spade, unlike mud, is *for* something. Because spades are essentially for digging, a good spade has whatever properties fit it for that end. Because mud does not have a function, it does not make sense to think of mud as being good or bad except in relation to some further function. This suggests as a working hypothesis the view that the goodness of x depends on it fulfilling the function of the kind K to which it belongs. This is a matter of degree and x will be better or worse to the extent that it fulfils that function. There are, however, some obvious concerns with this idea. In particular, it might be thought that the example of a spade has allowed me to make the case that the goodness of a thing depends on its function rather too easily. For spades (and cars) are unusual in being human artefacts, which have been *designed* for a purpose.

One possible concern is that a designer arbitrarily assigns a spade's function to it. As a result, saying that a spade has the function of digging fails to pick out any interesting metaphysical truth. We might just as easily assign it the function of being a doorstep or a weapon. Worse, it might then seem that a spade's function is irrelevant to questions concerning its goodness. The goodness of a spade depends entirely on the human interests it serves. A heavy spade with a sharp edge is good because it enables us to plant potatoes, bury treasure, and the like, not because it measures up to some abstract standard independent of human interests. The first point to make in response is that it is not in general a weakness of a theory that the presence of normativity and value in the world turns out to depend in some way on the exercise of rational agency. Realist views do not require that there would be normative truths in a lifeless universe comprised entirely of non-rational matter. Second, although it is true that artefacts such as spades are designed with the satisfaction of human interests in mind, we are still able to evaluate them independently of our *present* concerns. We can, for example, recognise the excellence, or shoddy workmanship, of a piece of obsolete or alien technology. So, even if artefacts are designed in response to human interests or concerns, it does not follow that the goodness of the artefact depends on satisfying any current interests. Should it be suggested

that in cases of obsolete or alien technology there is a covert reference to the designer's interests, we might reply that we are able to evaluate artefacts against the normative standard given by their function even if that function serves no possible human interest. Some weapons of mass destruction may fall into this category.

Perhaps, though, it will be suggested that it is the designer's intentions that matter. Certainly, it can be very difficult to know the function of an artefact if one is ignorant of the designer's intentions. However, it is not true that a designer can arbitrarily assign function to artefacts. One reason, as I have already mentioned, is that terms that refer to functional kinds, such as 'spade', are success terms. This comes from the fact that the function of an object depends on the criteria of identity for objects of that kind. So, a spade could be entirely unfitted for use as a doorstop or a weapon and still be a spade, but it could not be entirely unfitted for digging and still be a spade. Although a designer may well have a function in mind, he cannot decide that an artefact has a particular function by a simple act of will.

A second type of objection allows that the function of artefacts such as spades provide us with legitimate examples of goodness, but claims that artefacts such as spades are atypical insofar as they are the products of rational agency. Therefore, they can hardly serve as paradigm cases of goodness.<sup>14</sup> One possible response would be to argue that the world is the product of a Creator and therefore shot-through with rational design. A functional account of goodness seems to fit well with a theistic framework and it is probably not a coincidence that a number of prominent neo-Aristotelians are also theists.<sup>15</sup> Given certain

---

<sup>14</sup> Pigden (1990: 147–53) among others presses this line of objection against neo-Aristotelian accounts of goodness.

<sup>15</sup> I am thinking in particular of Anscombe, Geach and MacIntyre. The Unmoved Mover of the *Metaphysics* also plays an important and under-appreciated role in Aristotle's ethical theory. When in Book X Aristotle advocates a life structured towards the end of intellectual activity (*theoria*) he notes that it is 'not insofar as he is a human that he will live like this, but insofar as there is something divine in him' (1177b26-7). The relationship between a flourishing life and participation in the divine is brought out most explicitly in the *Eudemian Ethics*: 'If some choice or possession of natural goods – either goods of the body or money or of friends or the other goods – will most produce the speculation of god, that is best, and that is the finest limit; but whatever, whether through deficiency or excess, hinders the service and speculation of god, is bad.' (1249b17-20; see also 1145a6-11).

further assumptions concerning the attributes of God, this view would seem to have just the right shape for grounding the *normativity* of goodness. Although this is an important possibility and attractive in the present context, I shall simply acknowledge it and set it aside for the moment. As an atheist I do not think it is the right position to adopt and I am not persuaded that the theoretical advantages of theism in this case are sufficient to justify theism on balance. Although I shall return to related issues towards the end of the chapter, I do not see how to make constructive progress on the topic without too much digression here. In any case, I shall proceed without the assumption of a theistic framework.

A second possible response starts by observing that even if the world is not the product of rational design, we can still understand much of the biological world in design terms. In particular, evolutionary theory can explain a large number of apparently purposive features in nature. After all, evolutionary theory would not present a difficulty for the Argument from Design if it did not save (many) of the appearances. The case looks here particularly strong for organs that contribute to complex systems. For instance, we are inclined to say that hearts are for pumping blood, that eyes are for seeing, that roots are for absorbing nutrients etc. A plausible explanation for these biological functions can be given in terms of the evolutionary pressures bearing on an organism's ancestors.<sup>16</sup> The reason that the function of the human heart is to pump blood and not, say, to make a thumping sound is that only the former made a contribution to the fitness of our ancestors.

These examples fit well with a functional account of goodness. Given that the function of the eye is to see, it is very plausible that a good eye sees well. Perhaps more interesting, however, is the thought that evolution works at the level of behavioural traits. In particular, certain behavioural traits are adaptive for a species in the context of a distinctive way of life or 'life-form'.<sup>17</sup> For instance, climbing trees well is adaptive for squirrel monkeys because they primarily live on fruit that grows on trees. Signalling well is adaptive for

---

<sup>16</sup> The *locus classicus* for this view is Wright (1973). For a useful overview of evolutionary approaches to the nature of functions see Davies (2001: Ch. 2).

<sup>17</sup> See further Foot (2001: 33–7).

meerkats because they live in large social groups and are individually vulnerable to predators.<sup>18</sup> With such thoughts in mind, it looks attractive to say that the function or purpose of a particular behavioural trait should be understood in terms of its contribution to the way of life typical for the species. But, of course, man is also a species of animal with a distinctive way of life. Perhaps most significantly, we are as Alasdair MacIntyre (1999) calls us, ‘dependent rational animals’. We are also essentially social and linguistic beings. So, as with nonhuman animals, we might say that a behavioural trait is good when it is conducive to man’s distinctive way of life. But in the case of humans, a stable behavioural trait is just the sort of thing we call a virtue when it is good and a vice when it is bad. In this way, virtues and vices might appear indissolubly connected to our natural history.<sup>19</sup>

Perhaps Philippa Foot (2001) did most towards developing an account of ‘natural goodness’ along these lines.<sup>20</sup> The idea of natural goodness connects what is good for man with his natural history and his distinctive way of life in an attractive way. Presumably, we could have evolved quite differently, but we are as we are with our special strengths and vulnerabilities, and these facts have a strong bearing on what counts as living well for beings like us. It is also important to recognise that although evolutionary forces have produced design, or the illusion of design, an account of natural goodness can be constructed quite independently of evolutionary considerations. What matters is the place of particular traits in the life of a species *as they are*. What Michael Thompson calls ‘natural-historical judgments’ (2008: 20) concern the present nature of a species. These have an evolutionary explanation, but the norms consequent upon

---

<sup>18</sup> Looking back at one of earlier examples, we might similarly talk about the goodness or badness of an apple in terms of its contribution to the reproductive cycle of the apple tree. I mentioned the following criteria above: having a pleasing taste, being pest-free, being relatively blemish-free and storing well. We can now see that the first two are good because an apple that does not have a pleasing taste and is not pest-free will not make its proper contribution to the life-form of an apple tree. The criteria of being relatively blemish-free and storing well are perhaps less clear and remind us that we can evaluate an object according to its intrinsic goodness or from an external perspective. For instance, it is certainly good for a mouse to be quick and lean given its way of life. From a cat’s perspective, however, it is better for a mouse to be slow and fat. I return to this topic towards the end of section 6.4.

<sup>19</sup> As Geach wrote, ‘men need virtues as bees need stings’ (1977: 17).

<sup>20</sup> Recent discussions include Thomson (2008), Thompson (2008) and Hacker-Wright (2009).



our natural history do not depend on an evolutionary justification.<sup>21</sup> This may help advocates of an ethics of natural goodness avoid the charge that they are trafficking in explanatory reasons under the guise of justificatory reasons.<sup>22</sup> What it is to be a good man depends on what it is to be a man, but not on the causal or historical explanation of the nature of our species.

I think there is a lot to be said in favour of this type of view and much to be learnt from it concerning the nature of goodness. For instance, I am inclined to think that it provides us with the resources to say most or all of what we should say concerning the goodness of most or all nonhuman life. Ultimately, though, I am not convinced that it is adequate as it stands as an account of human goodness. For, to misappropriate a phrase, it is distinctive of human rationality to put nature to the question. Whilst one can take a more or less positive view of the natural mode of human existence, it makes sense to adopt an external perspective on our distinctive way of living and ask whether it provides us with an ideal to be pursued. When one does this, it seems that there is always the possibility that we will not reflectively endorse the way of life distinctive of our species. For instance, it would be hard to deny that human beings are warlike animals or that aggression between sub-groups of humans has been a pervasive and non-incidental feature of human life and history.<sup>23</sup> It would be curious if an anthropologist Mars omitted these facts from his report on us. However, the conclusion that warlike behaviour is good for human beings does not follow from this.<sup>24</sup> I am sure that Foot and other advocates of natural goodness would agree and I think the most likely response would be that aggression and war are harmful given the nature of man. Men do not flourish under conditions of warfare. Although this is right, it is not clear how this can be related back to the idea of natural goodness. If warfare is a part of our distinctive way of life, it appears that an independent criterion of goodness is being employed.

---

<sup>21</sup> Compare with my earlier comments on the relationship between a designer and a spade's function.

<sup>22</sup> See, however, Millum (2006) for the complaint that this constitutes a move away from the relatively well-understood idea of evolutionary function that partly motivates the idea of natural goodness in the first place.

<sup>23</sup> Slote (2003: 134) presses this objection in his review of Foot's (2001).

<sup>24</sup> Similarly Glassen (1957) argued that Aristotle's 'function argument' conflated what it is to be a good man with the good *for* man.

Foot comes close to addressing this concern when she asks what reason there is to care about natural goodness (2001: 52f.). Her response is essentially that the pursuit of the human good is the rational choice. She supplements this with the argument that there is a closer connection between happiness and natural goodness than is generally recognised (2001: 94–7). This might seem to miss the point. For the question I was pressing above is not whether we have reason to pursue the human good, but whether we have reason to act in accordance with norms of natural goodness. But, in fact, Foot’s response does address the question. For, if the human good is natural goodness, then the question of whether we can reflectively endorse natural goodness just is the question of whether we can reflectively endorse the human good. The problem, of course, is that the identity claim is precisely what is at stake and so the work Foot puts in to persuading us that we have reason to care about natural goodness is wasted if we are not already persuaded of her basic position.<sup>25</sup> Still, the fact that the question of whether we can reflectively endorse natural goodness *appears* distinct from the question of whether we can reflectively endorse the human good is *prima facie* evidence that natural goodness is not the whole story. Indeed, I shall argue below that it is not.

### 6.3 Goodness and the Practice of Chess

This may be an appropriate juncture at which to recap and take stock. I started with Geach’s idea that ‘good’ typically functions as an attributive adjective. It normally makes no sense to say that ‘x is good’ unless some kind K to which x belongs is understood. After looking at the metaphysical implications of this, I endorsed Urmson’s view that ‘good’ is typically a term used to evaluate things against an appropriate standard for the kind of thing being evaluated. In particular, I picked up the Aristotelian thought that what provides the standard against which the goodness of a thing is measured is the function of the kind to which it belongs. In response to the anticipated objection that most things are unlike artefacts and do not have functions, I considered the idea that evolutionary processes produce a category of natural goodness that applies to the biological

---

<sup>25</sup> This is not of course a criticism of Foot’s approach. One can hardly move forward on the basis that one’s earlier arguments are wrong.

world including man. Given that there is something that it is to be a man, there is something that there is to be a good man. Finally, though, I stopped short of endorsing Foot's account of natural goodness with the observation that there appears to be a further question of whether we can reflectively endorse the type of natural goodness given by our natural history.

I would like to bring forward five ideas from the above discussion. First, in the central case, the goodness of a thing depends on the kind of thing it is. Second, evaluating the goodness of a thing is a matter of determining the extent to which it measures up to a normative standard. Third, the appropriate normative standard for a kind is teleological in the sense that it depends on the function or purpose of a kind. Fourth, the function or purpose of a kind depends on the criteria of identity for a kind, i.e. what makes it what it is and not another thing. Fifth, given something's goodness, the question remains open as to whether we should reflectively endorse the pursuit of it. In this section I shall bring the first four ideas to bear on the toy example of goodness in chess. In section 6.4, I shall look at deontological constraints in chess. This will give us a sketch of the good and the right within the context of a particular practice. I shall then be in a position to pull focus and consider the cases of morality and epistemic normativity in the final section of the chapter. In doing so I will return to the fifth idea above. To anticipate, I shall be arguing that the fifth idea is accounted for by a second sense of what it is for something to be good and that this is related to questions of meaning in life.

At the beginning of the chapter I advertised that I would offer an account of goodness in chess. Rather than asking in general terms what goodness in chess amounts to, however, I would like to start with the more tractable question of what it means to be a good chess move. This meets the demand that attributions of goodness require a kind, i.e. chess move, and allows me to build up an account from the bottom. So, what is the appropriate normative standard against which to measure a chess move? Well, given that the normative standard for a kind depends on the nature of the kind in question, ascriptions of goodness or badness to chess moves must depend on some 'chessy' feature. It also seems that the chessy feature must be related to the practice of which a particular move is a

part. Altering the location of a horse-shaped piece of wood would not count as playing a knight move outside of the practice of chess. Moreover, given our earlier considerations, we should look for a normative standard related to the function or purpose of a chess move. I propose, then, that a good chess move is one that is conducive to the goal of the practice of chess. To a first approximation, this means a move conducive to securing the best possible result in a game. This could be a move conducive to winning a game, but it could come to mean achieving a draw from an inferior position or maximising the chances for the best possible result, whilst not necessarily leading there with best play.

It might seem that this idea of goodness in chess faces the same difficulties as the instrumentalist account of epistemic normativity that I criticised in the last chapter. If the value of an action derives from the fulfilment of some further end, why not adopt some other policy to achieve that end? Why not, for instance, cheat? In fact, however, I am not proposing that we think of the goodness of a chess move purely in instrumental terms. Rather the relationship between good moves and a winning game of chess is partly to be explained in constitutive terms. So while it is true that good moves are instrumental to winning a game of chess, it is also the case that a winning game just *is* a series of good moves. The model I have in mind is Aristotle's account of flourishing, according to which it is part of the definition of virtuous action that it is conducive to flourishing, but also the case that flourishing is constituted by virtuous action. To lead a flourishing life is not something over and above acting virtuously. If the goodness of a chess move were to be explained in purely instrumental terms, then spiking your opponent's water with sleeping pills, say, would turn out to be a good thing to do from a narrowly chess-focussed point of view. It would help to win the game. But sabotaging one's opponent does not count as a good from any perspective, because such an action, aside from its moral viciousness, is no part of what counts as winning. Spiking drinks with sleeping pills would have no proper place in the description of a winning game.

Another way to bring out this constitutive connection between good moves and the aims of chess is to reflect on possible variations of Moore's paradox. There

is something troubling about the assertions ‘ $p$  is true, but I do not believe that  $p$ ’ and ‘ $p$  is true, but I believe that not  $p$ ’.<sup>26</sup> The problem is not, of course, a logical contradiction. The truth conditions for these assertions are straightforward, and the assertions could be true. It is only first person assertions in the present and future tenses that are troubling. One plausible explanation of this is that there is a norm of truth governing belief.<sup>27</sup> If I sincerely assert that  $p$  is true, but do not believe  $p$ , or believe not  $p$ , then I violate this norm. Given that a norm of truth governs belief, it is irrational not to believe a proposition I take to be true. Notice, though that there would be no incoherence or conceptual confusion if belief were governed by the norm of truth in a purely instrumental fashion. For then the norm I would violate in failing to adjust my beliefs to what I take to be true would be the norm of truth-conduciveness. But it is easy to think of examples in which failing to adjust my beliefs to what I take to be true would actually be conducive to truth. It is always possible to arrive at true beliefs on the basis of false ones (e.g. Gettier cases) and in these instances it would be better not to adjust my beliefs to what I take to be true. Similarly, it would be better not to adjust my beliefs to what I take to be true when I take a proposition to be true on the basis of misleading evidence. So, if the relationship between belief-formation and the end of truth were instrumental, there would be cases in which assertions of the form ‘ $p$  is true, but I do not believe that  $p$ ’ would violate no norm in relation to the end of truth. But, in fact, first person present tense utterances of Moorean sentences are always incoherent. It is difficult even to make sense of the idea that a rational agent could fail to adjust their beliefs to what they take to be true. This can be explained, I suggest, by the fact that the regulative role truth plays with regard to belief is constitutive rather than instrumental.<sup>28</sup>

---

<sup>26</sup> There is a connection between the strangeness of asserting ‘ $p$ , but I do not believe that  $p$ ’, and my claim in the previous chapter that it is irrational to fail to form the belief that  $p$  when one encounters overwhelming evidence in its favour. The connection is revealed by Evans’ idea of transparency, according to which an agent answers the question of whether she believes that  $p$  in the same way as she answers the question of whether it is true that  $p$ , namely by attending to the world rather than her own mental states (1982: 225–6).

<sup>27</sup> I return to this idea in the next chapter.

<sup>28</sup> See Railton (1994: 71–3) for related discussion of Moore’s paradox and truth as the constitutive aim of belief.

If my proposal that there is a constitutive relation between good chess moves and the goal of chess is right, similar considerations should apply. In particular, there should be something deeply confused about the assertion ‘moving thus leads to winning the game, but I shan’t do it’. And, indeed, this assertion *does* seem odder than the assertion ‘spiking my opponent’s water with sleeping pills leads to winning the game, but I shan’t do it’. The first assertion stands in need of further explanation in a way that the second assertion does not. Spiking an opponent’s water is no *part* of winning a game of chess and so the fact that it leads to that end does not give us a reason to do it. This is some confirmation that we were right to say that the goodness of a chess move is not merely its instrumental goodness in relation to winning the game.

But this comparative oddness may seem scant consolation for the fact that the assertion ‘moving thus leads to winning the game, but I shan’t do it’ is not incoherent in anything like the same way as the Moorean assertion ‘*p*, but I don’t believe that *p*’. I think, however, that there is a straightforward explanation at hand. For when we are considering what makes sense within the context of a game of chess, we naturally bring other normative standards to bear. We can think of plenty of good reasons why we might recognise a winning move and decide not to play it. Perhaps we recognise that our opponent is psychologically vulnerable in such a way that losing would be harmful to him. Perhaps he will sink into a deep depression. Or perhaps we have some financial incentive not to win. Perhaps we are subject to intimidation from the ruling authorities. Given these possibilities, we naturally assume that if someone asserts ‘moving thus leads to winning the game, but I shan’t do it’ they are responding to the all-things-considered judgment that they have most reason not to win. We would be curious for a further explanation, but we would not be baffled into incomprehension. By contrast, we might try to imagine a world in which playing chess constituted the only activity. I suggest that the inhabitants of this impoverished world would find the sincere assertion ‘moving thus wins the game, but I shan’t do it’ similarly baffling to Moorean assertions for us. For, if chess were the only activity, there could be no rational reason not to pursue the good of the practice.

So, it is not difficult to provide an account of the way in which reasons that are external to the practice of chess block the practical inference from ‘moving thus wins the game’ to ‘I shall move thus’. It is interesting, however, to consider a different sort of reason for failing to conform to the norm of winning at chess. For example, it is possible that we want to encourage a learner to develop his talents and, so, stay our hand. We might think that it makes no sense to ‘punish’ a learner for an error that he could not reasonably be expected to have avoided.<sup>29</sup> This sort of reason is connected to the wider practice of chess and shows the inadequacy of my earlier conception of chess according to which the goal of participation in the practice is to win games. Chess has institutions, history, traditions, literature, physical artefacts, tournaments and much else associated with it. If it did not have at least some of these features it would not count as a practice or a rule-governed activity (and, simultaneously, part of what makes these disparate elements part of a single practice is their contribution to the goals of the practice). Training junior players is a good internal to the practice. Indeed, it is necessary for the survival of the practice. So, failing to make the best move in order to train an apprentice player can be justified in terms of chess goods. What we should really claim, then, is that it is incoherent, from a perspective internal to the practice, to assert ‘doing thus is conducive to the flourishing of the practice of chess, but I shan’t do thus’.

In some ways this would be a much better account. Certainly, it is a better analogue of the ethical case, where much of the work of saying what counts as a virtue depends on the work of judging what it is to flourish. Unfortunately, however, what it is for the practice of chess to flourish is far less clear than what it is to succeed in playing a particular game and so a fuller account tends to obscure the basic normative structure. One reason that chess is sometimes used as a toy example for discussions of normativity is that it appears to be a tightly

---

<sup>29</sup> Similar considerations apply to the practice of philosophy. Although philosophy is governed by the norm of truth, it takes experience and sensitivity to know which of a student’s false beliefs to correct and when. We might also notice that the normative standards for good philosophy seem to move as the student develops. A good undergraduate essay would normally make a bad doctoral thesis, and I daresay that a good doctoral thesis is not normally a good work of philosophy *tout court*. But this does not really show that philosophy has moving or multiple standards of evaluation. Rather it shows that to evaluate someone’s work is often just a way of evaluating a person at a particular stage of her development.

rule-governed activity. This may be true of a particular game, but a game is only comprehensible within the context of the practice, the boundaries of which are far less clear.<sup>30</sup> Indeed, there is a general problem of establishing the identity conditions for practices. This difficulty is even more pressing in light of the account of deontological concepts I develop in the next section. Unfortunately, I do not see a general, principled way of solving the difficulty. Of course, it is possible to place strict conditions on which activities count as instances of a practice. We could *define* chess in such a way as to exclude near variants, but this would be an arbitrary and unhelpful move. First, it would have little to do with carving nature at its joints. Second, the criteria of identity for practices such as chess are necessarily indeterminate in order to allow for innovation and development of the practice.<sup>31</sup> This is further complicated by the fact that some practices have more rigid identity conditions than others. Compare, for example, Byzantine iconography and abstract expressionism.

Although these are genuine difficulties, I do not think they are anything like fatal to the position I am defending, namely that the goodness of a chess move is dependent on its role, or function, within a practice and, more particularly, dependent on whether it contributes to satisfying the goal of the practice (which partly determines the identity of the practice). Although it should be allowed that the identity conditions for practices are sometimes vague, this point should not be overstated. We might argue whether *Waiting for Godot* is a tragedy or a comedy, or more one than the other, but we know it is a play and not a painting. Similarly, we are perfectly able to distinguish central instances of chess from central instances of squash or tiddlywinks. Doubts concerning what should count as an instance of a practice occur mainly in relation to marginal cases and

---

<sup>30</sup> In fact, the rules governing a game of chess are far less clear than it might seem. For instance, positions sometimes occur in which there is no prospect of either player winning without one player making an obviously bad move. Normally players agree a draw in these circumstances. However, chess is played against the clock. Each player has a fixed length of time to play a certain number of moves and if he runs out of time he loses. So, if one player has more time than his opponent, one possible strategy is to keep playing moves in a drawn position until his opponent loses on time. To prevent this, if a player has less than two minutes remaining on his clock, he can appeal to the tournament arbiter and object that his opponent is 'making no attempt to win by normal means'. If the arbiter agrees, a draw will be awarded. Although it is not normally controversial, the arbiter's decision depends on practical wisdom rather than the application of formal rules. He is required to read a player's intentions.

<sup>31</sup> See further van Fraassen (2002: 145) on the nature of scientific practice.



in response to two types of considerations. First, it can be uncertain whether the norms internal to a practice apply to particular new cases – “it’s all very well, but is it Art?” Second, though, there can be challenges to the norms internal to practices – “isn’t there more to Art than sweetness and light?” These are difficult normative questions, but we regularly address them as practices evolve and change in response to external norms and pressures of internal consistency. We shall also see in the final section of the chapter that these questions do not arise in quite the same way in relation to ethical and epistemological questions. I propose, then, to bracket them off for the moment.

#### *6.4 Rightness and the Practice of Chess*

It is customary to draw a distinction between the good and the right. Sometimes philosophers also propose to explain the one in terms of the other. For instance, Moore claimed that ‘right’ simply means ‘productive of the greatest good’ (*PE* 18). On the surface of things, however, this looks unlikely to be true. It is wrong as a purported lexicographical fact and it is implausible as a claim about the hidden conceptual structure of normative discourse. For, there are a great many apparent differences between the good and the right. We have seen that goodness is normally understood in relation to an end, although the connection is not wholly instrumental. By contrast, to say that an action is wrong is normally to say categorically that one must not act thus. When we say that something is the right thing to do, we sometimes mean that is the right thing to do *despite* its consequences. It is also true that we sometimes say that an action cannot be right in view of its consequences, but nonetheless we think that the two are logically distinct.

There is also a difference in the normative force associated with the good and the right. The good is weaker. Doing what is good often seems optional – “It would be good/better if you did it”. Whereas we can oppose the goodness of an action with other considerations, a claim that an action is right is a conversation stopper. Doing what is right is doing what one’s duty dictates, what one must do or what one is obliged to do. By contrast, the connection between goodness and obligation is weak. ‘Ought’ implies ‘can’, but ‘good’ does not imply ‘can’. So, it would be good if I could run a mile in less than four minutes. But I am under

no obligation to be able to do this. Related to this, the connection between goodness and blame is weak. All things being equal, I am not blameworthy because I cannot run a mile in less than four miles, but I am blameworthy if I fail to fulfil my obligations.

In this section, I shall offer an account of the place of the right within chess that is consistent with these differences. It is complementary to my account of the good and by the end of this section we shall have a complete sketch of the norms of chess. Under what circumstances, then, might we distinguish between the right and the good in chess? One example is strategy. It is not strictly speaking wrong to open a game by moving an outside pawn as beginners sometimes do. However, it is bad in the sense that it wastes time in developing pieces, cedes the centre of the board and weakens one's pawn structure. Therefore, it is not conducive to winning the game and a bad move. But, one *may* open thus and a beginner does not merit any sanction if he does so. This fits with our earlier description of goodness in chess.

By contrast, it is wrong to break the rules of the games, for example, by moving a rook diagonally or by using pawns as missiles to bombard the opponent's king into submission. This is not because such moves are bad ones. Rather it is because these are not chess moves at all. Indeed such 'moves' undermine the practice and thus the conditions under which goodness in chess is possible. In order to play chess it is necessary that we follow the rules that serve in part to define the game. If a player 'checkmates' his opponent, but has two bishops on squares of the same colour at the end of the game (and has not promoted a pawn to a bishop), then he has not won. It does not make a difference if his opponent has not noticed the rule violation and concedes the loss. He may be awarded the win and its trappings, but all that shows is that everyone is persuaded of a falsehood. It is just as if someone were presented a gold medal for winning a marathon when they had travelled half the distance by taxi. They cannot have won the marathon for the simple reason that they have not run a marathon. So, insofar as we are playing chess at all, we must obey the rules. Failing to do so fundamentally undermines the activity and also bars one from any goods which are internal to the practice. The problem is not merely that breaking the rules is

unconducive to attaining the good associated with the ends of the practice, but that breaking the rules makes it impossible to act well or attain those goods. For goodness comes with the game. It is not a property that can be predicated of a move independently of its role in the game.

One might recall here my earlier observation that terms referring to functional kinds are success terms. For instance, I suggested that there could be no such thing as a spade made out of room temperature blancmange. The thought might then occur that there is a connection between the good and the right in the sense that something can be such a bad instance of a kind that it no longer counts as an instance of that kind at all. We can imagine, for example, a good spade that is left to rust. The spade will gradually deteriorate and become worse. Finally, it will become more rust than spade and so bad that it will no longer count as a spade at all. This may suggest that the difference between the right and the good is less clear than I have suggested and that the good and the right (or the bad and the wrong) lie on a continuum. For, by analogy, it might seem possible to play chess so badly that one is not playing chess at all. One might think of a situation in which a player makes legal moves at random, or deliberately makes bad moves.

In these sorts of cases I am inclined to agree that the player would not be playing chess. However, it is not the badness of the moves that matter but the mental states of the player. A beginner can play very badly indeed without there being any doubt that she is playing the game. Indeed, complete beginners often lose in just four moves to ‘scholar’s mate’.<sup>32</sup> But as the term ‘scholar’s mate’ indicates, its victim is normally a very junior practitioner. If an experienced player lost in four moves, we would know that they were not really playing the game, but using the game as a vehicle for some other end. The important point here is that participation in a practice depends on the intentions of the practitioners. All things being equal, an experienced player could only lose in four moves if his intentions were unrelated to the norms that govern chess. Although one can

---

<sup>32</sup> It is possible to lose in just two moves to ‘fool’s mate’ (1. f3 e5 (or e6), 2. g4 Qh5 ++). However, this requires a particular sequence of moves from both players and is more likely to befall a stooge than a fool.

intend to play chess and fail to do so because one has not grasped the rules, one cannot play chess in spite of one's intentions or in the absence of any intention to do so. For example, one cannot find oneself playing chess by accident. This explains a number of other situations in which performing a series of legal moves would not count as playing a game. For instance, one could pretend to play chess, say on a stage, and make a series of legal moves by accident. Or, one could move the pieces legally, but as someone else's agent. In neither scenario would one be playing chess.

So, conforming to the rules is a necessary condition, but not a sufficient condition, for participation in the practice of chess. Another necessary condition is that playing chess requires intentional action governed by the norms of the game. Playing chess is something one *does* well or badly. It is a manifestation of rational agency. Only in a derivative sense can we talk of *being* a good chess-player or a good move. In this sense, chess moves are unlike spades. Chess computers provide an interesting point of comparison here. For we can describe them as 'good' or 'strong' chess players and, indeed, they can compete with the best players in the world. However, chess computers do not (currently) *do* anything in the relevant sense, for they are not (yet) capable of intentional action. Consequently, I suggest that a computer no more plays chess than an actor on a stage plays chess, or a monkey at a typewriter composes *Hamlet*, or Searle's unfortunate captive speaks Chinese. They are not playing chess because they are neither responsive to nor accountable to the norms of the practice. Or, as it sometimes put, they are conforming to the rules but not following them. As a result, computers merely *simulate* playing chess. Like spades, we can evaluate chess computers in light of their function and independently of human interests. We can 'play' two computer programmes against each other and get a reliable measure of which one is better. But, like spades, we cannot talk meaningfully about the duty, obligation, rightness or wrongness of a chess computer.

So, wrongness in chess is a matter of intentional actions that undermine the preconditions for engaging in the practice. What, though, of the normative force I have already mentioned as being associated with deontological concepts such as the right? It seems there may be a clue in the etymology. The term

‘deontology’ derives from the Greek *deon*, which is often translated as ‘duty’. However, the word ‘duty’ is a relatively modern, semi-technical term of art correlated in particular ways with rights and obligations. In fact, there was no ancient Greek word for ‘duty’. As Bernard Williams (1985: 16) notes, deontology comes from the Greek for what one *must* do. So, etymology suggests a connection between deontology and necessity. Indeed, this coheres with what we have already said, namely that obeying the rules of chess is a *necessary* precondition for engaging in the practice. Another way to put this is to say that we *need* to obey the rules if we are to play chess. I should like, then, to conclude my discussion of rightness in chess with a few comments on the subject of needs. I shall develop the relation between needs and deontology further in the next section.

To say that there is a need for something is often to make a particularly strong sort of claim. As Harry Frankfurt puts it, ‘claims based upon what a person needs have a distinctive poignancy’ (1984: 1). But, of course, not all needs have the same poignancy. It is part of the grammar of the term ‘need’ that one needs *x* for *y* or *in order to z*.<sup>33</sup> The poignancy of a particular need is only revealed when the placeholders have been filled. We can then distinguish between the needs we need to have and the needs we do not need to have. The needs we do not need to have are contingent upon our desires and interests.<sup>34</sup> For instance, I might need £30 to buy a ticket to watch a football match. But if I tried to press this needs claim on a third party, it might well be responded that I do not really need the £30 because I do not really need to watch a football match. What matters when we are assessing the strength of contingent needs is the value of the end. There is no general moral requirement to satisfy anyone’s contingent needs and, if the end is nefarious, there may be a moral requirement to prevent the need being satisfied. Nevertheless, contingent needs can have something near the force of moral obligation if the value of the end is particularly high, the need is particularly easy to satisfy or if one stands in a morally significant relationship

---

<sup>33</sup> Because of the relationship between the satisfaction of needs and action, it may be better to restrict needs in the reason-giving sense to ‘needs to’ as opposed to ‘needs for’. I thank Jonathan Lowe for drawing my attention to this point. See also his (2006) in this connection.

<sup>34</sup> They are variously called ‘volitional’ (Frankfurt 1984), ‘contingent’ (Reader and Brock 2004), or ‘instrumental’ (Wiggins 1998) needs.

with the bearer of the contingent needs. It is clear that when I have been discussing the necessary preconditions for playing chess I have been discussing contingent needs. We may need to obey the rules if we are to play chess, but we do not need to play chess. The strict necessity of obeying the rules only binds us insofar as we are participants in the practice.

### *6.5 Non-Contingent Needs, Flourishing and Meaning*

In the previous two sections, I sketched an account of the right and the good in the context of the practice of chess. Briefly, I argued that the goodness of a chess move is a matter of its conduciveness to the goal of winning. This goal is the regulative norm for chess moves and the relationship between the two should be understood in constitutive rather than instrumental terms. I then argued that the right is associated with the necessary preconditions for playing chess and that these are contingent needs. Providing an account of value and obligation within the context of a particular practice is not a trivial task. However, my real quarry lies elsewhere and I would now like to consider how, and to what extent, the normative framework I have sketched for chess can be applied to moral and epistemic normativity.

We have already noted that the normative significance of the reasons internal to the practice of chess can be questioned. The reason that a sentence like ‘moving thus leads to winning the game, but I shan’t do it’ doesn’t strike us as paradoxical is that we can bring non-chess related norms to bear on the practice. We can consider how the practice of chess fits into a person’s life and we can consider its significance in terms of more fundamental values. We can ask whether chess contributes towards or undermines a person’s other projects. Hence, we can allow that an action is good relative to the norms internal to the practice of chess, but ask whether it is good in an all-things-considered sense or in a moral sense. There are certainly some practices which are morally evil, or indifferent, or largely good but liable to come into conflict with moral demands. The point is that we can evaluate chess from an external perspective. If we decide that playing chess is not a good thing in light of our other values, we can decide that we should stop playing.

I suggested above that the reason some needs are contingent is that they do not need to be met. Thus, chess related needs are contingent because we do not need to play chess. By contrast non-contingent needs are needs that do need to be met. But, what is it to say that a need needs to be met? The idea is not merely that it would be good or valuable for the need to be met, but that they have the binding force of obligations. Soran Reader and Gillian Brock (2004) make a useful proposal here. They suggest that we should mark the distinction between contingent and non-contingent needs as follows:

The mark of moral importance of non-contingent needs in ethics is that the needing being simply cannot go on unless its need is met. It is no exaggeration to say that in a state of non-contingent need, the very existence of the needing being as we know it is at stake. (2004: 252)

They go on to mention the examples of ‘agency, life, flourishing or avoidance of harm’ (2004: 252) as candidates for non-contingent needs and argue that they generate obligations at the level of public policy. I think that they are right to stress the connection between non-contingent needs and the persistence of the needs-bearer. I also think they are right that when the persistence of a needs-bearer is at stake, we have a moral obligation to meet his needs. Building on these insights, I would like to look more closely at how non-contingent needs should be understood within the normative framework I have been developing.

I suggested above that the normative force of the claim ‘S needs x in order to y’ is revealed when the placeholders are filled. Although it is a little awkward grammatically, we might say that non-contingent needs are characterised by true propositions of the form ‘S needs x in order to be S’. So, ascriptions of non-contingent needs depend on questions of personal identity. We cannot say whether a person has a non-contingent need unless we have a grasp of what it is for a person to persist and we cannot say what it is for a person to persist without a grasp of what it is to be a person. This might remind us of the earlier thought that the appropriate normative standard for a kind depends on the criteria of identity for a kind. Moreover, we have already seen that the right in chess depends on the criteria of identity for the practice. Let us then rephrase part of the conclusion of section 6.4 as follows: ‘A chess player needs to obey the rules in order to be a chess player’. So, the necessary preconditions for participating

in the practice of chess generate non-contingent needs *within* the context of chess. Were a subject's agency restricted entirely to playing chess, chess related needs would be non-contingent. Breaking the rules of chess would undermine the preconditions for the only sort of value available to the subject. However, we can ask whether a person needs to be a chess-player. As the answer is almost certain to be 'no', the need expressed by the sentence 'a chess-player needs to obey the rules in order to be a chess player' is a contingent need.

My proposal, then, is that non-contingent needs are categorical obligations and that they depend on what is essential to the nature of the needs-bearer. At this point it would be desirable to have an account of the criteria of identity for persons; what it is that makes a person a person and not some other thing. Unfortunately, I do not have a detailed account to offer. Nevertheless, I would like to venture some thoughts on the general shape of such an account. An obvious starting point is that we are embodied beings and, hence, physically vulnerable. As a result we have so-called 'basic needs'. These are goods such as food, water and adequate shelter that are needed to hold body and soul together. As basic needs are shared by everyone and do not depend on a controversial account of personhood, they have figured large in public policy debates.

Basic needs are non-contingent needs. However, they provide rather a thin basis for normativity. It was with something like basic needs in mind that Shakespeare had King Lear say:

O, reason not the need! Our basest beggars  
Are in the poorest thing superfluous.  
Allow not nature more than nature needs,  
Man's life is as cheap as beast's. (*King Lear* II. 4)

Of course, the fact that a person's life can go very badly despite their basic needs being met does not show that our obligations go beyond meeting those needs. Nevertheless, it is clear that our persistence depends on more than biological continuity. Most significantly in this context, man is an essentially rational animal. By this I mean that it is a necessary part of our identity that we recognise reasons for action and belief and that we are capable of responding to those reasons by modifying our behaviour and thought. This has important



implications. For example, the capacity for reason requires that we can represent the world as being in certain ways. It also requires the capacity to reflect upon these representations, which in turn depends on an ability to construct sentences that contain or refer to other sentences used to express our reasons. Moreover, we are social or political beings. We are born dependent upon others and that is normally how we live our lives and how we die. We learn language from others, and if language is necessary for reason, then so is company. As well as being essentially rational, language-using, social beings, we each have a practical identity. We are mothers, women, Europeans, teachers, students, plumbers, lovers, violinists, burglars, Muslims, Tories, vegans and the like. The components of our practical identity are contingent, but the fact that we have a practical identity is not.

I submit, then, that non-contingent needs are the necessary preconditions for the persistence of beings like us in an environment like ours. Furthermore, it is part of the human condition that we are rational agents and this fact places obligations upon ourselves and upon others. We have no choice but to act and we cannot escape the fact that our actions have moral significance. That is our plight. For, being rational, we are accountable to reasons. Moreover, we have no choice but to represent the world and to form beliefs. We can choose whether we play chess, but we cannot choose whether to be moral or epistemic agents. Morally significant action and belief-formation are non-contingent activities.

Despite this crucial difference between activities that are essential to our nature and contingent practices, the normative structures governing them have important similarities. I associated the rules for playing chess with the necessary preconditions for engaging in the practice. Similarly, there are necessary preconditions for moral and epistemic agency. Whereas the preconditions for chess are contingent needs, the preconditions for moral and epistemic agency are non-contingent needs. The necessary preconditions for moral agency include the capacity for rational, autonomous action. The necessary preconditions for epistemic agency include conformity to the basic canons of logic. If these preconditions are undermined, there comes a point when we cannot go on as rational agents. This places obligations upon us that must be met in our own

case and in the case of other rational agents. For, it is a basic principle of rationality that we treat like alike. It is also a metaphysical presupposition of my view that we are each one among many equally real people.

So, meeting the necessary preconditions for moral and epistemic agency situates us in the realm of value and opens up the possibility of goodness. It also enables us to act badly and it is a consequence of my position that there can be tension between the right and the good. As I discussed earlier, one must be engaged in the practice of chess to be a victim of scholar's mate. If this is right, we also have a *prima facie* moral obligation to preserve a life even when it is going badly.<sup>35</sup> Given the possibility of acting well or badly, there is also the possibility of acting more or less well. As I argued above, playing chess well is a matter of degree and depends on meeting the normative standards for the practice. In the simplified case of a single, competitive game this means playing moves that are conducive to, or increase one's chances of winning. Similarly, there is a purpose to epistemic activity and that provides a normative standard against which belief formation can be measured. In the next chapter I shall defend the view that the goal of belief is truth. I shall also explain how this relates to questions of knowledge and justification and, so, provide an account of what goodness in the way of belief amounts to.

Although morally significant action is like belief-formation in the sense that it is inescapable it is more difficult to account for the moral good because morality has no uncontroversial goal. Rather than defending a substantial account of the moral good, however, I shall be content to outline the structural features of moral normativity. To this end, I propose to draw a distinction between morality and ethics. Morality, as I have argued, is a matter of meeting the necessary preconditions for engaging in non-contingent activities; activities that beings like

---

<sup>35</sup> The *prima facie* obligation to preserve life does not entail that lives should be preserved in all circumstances and all costs. One very general reason, as I discuss below, is that we can always turn our back on the realm of normativity. A second reason is that death is certain for beings like us. As I have linked non-contingent needs to our persistence conditions, it might be objected that I am committed to an obligation to be immortal. In fact, as death approaches there is less that *can* preserve our identity and so our needs diminish. For this reason, I am also inclined to say that non-contingent needs must be stable or enduring, although I do not know of a principled way of deciding *how* stable or enduring.

us to cannot but engage in. By contrast, I propose that ethics is best understood as an activity of practical reason with the goal of ordering the field of practices towards the end of flourishing.<sup>36</sup> For once our non-contingent needs are met, we are left with a plurality of sources of goodness associated with a plurality of practices. The form of practical reason that is distinctive of ethical thought allows us to put them together, so we can answer the Socratic question, “how ought one to live?” So, once we have discharged our moral obligations, ethics is a matter of bringing the various sources of goodness available to us into a balanced and coherent life. For example, it is a matter of deciding whether to structure a life in pursuit of a single overarching aim, or to pursue a number of goals.

We have, then, a sketch of the right and the good. Although it is little more than a promissory note, the broad outlines should be clear. Within the category of value-makers, I have identified separate right-makers and good-makers associated with the necessary preconditions and the goals of practices respectively. It might be considered a weakly naturalistic account of normativity in the sense that it does not depend on supernatural agency or otherwise spooky entities. But whilst my ontology has not outstripped the ontology of the natural sciences, or not obviously, I have relied on distinctively philosophical methods such as conceptual analysis, thought experiments and *a priori* metaphysics. In this sense, it is a non-naturalistic account. I have also appealed to our experience as agents situated within practices. This involves a practical commitment to the norms internal to a practice. For instance, we do not, *whilst we are playing chess*, doubt certain normative truths. If we notice a combination of moves which forces a win in short order, we do not doubt that it is a good strategy. When we are playing and have grasped the rules that govern the practice, we do not wonder whether it might not better to move our pawns sideways or diagonally. By contrast, a spectator who does not know how to play might wonder such things. Or knowing the rules and the goal, she might wonder

---

<sup>36</sup> I use the term ‘flourishing’ as a placeholder for the well-lived life, the good life or the objectively happy life and I take it to be consistent with a wide range of normative theories. Given the place of flourishing within my overall theory, however, it is not blandly neutral. For, a flourishing life cannot be constituted by activities that undermine the preconditions of moral agency. Moreover, it must be associated with the exercise of the capacities that are essential to our identity.

whether the game might not be better were pawns permitted to move sideways. Distancing herself from the flow of the activity and adopting an external standpoint, a player might also ask such questions. In doing so, she abandons her practical commitments and adopts a stance towards chess more akin to that of a scientist's. It is at this point that the Humean doubts we encountered in chapter one can creep in and one can start to doubt the normative force of the sorts of facts I have identified as value-makers. Normative reasons are supposedly queer, it will be recalled, due to the phenomenon of normative force. Part of the explanation, then, of why normative reasons are difficult to account for using only the methods of the natural sciences is that when an agent is not practically committed to an activity, there is no necessary connection between the normative force of the right and the good on the one hand, and her motivational mental states on the other. Rather than pursuing this subject here, I shall devote chapter eight to questions of motivation and the relationship between normative force and desire.

### *6.6 Life's Meaning and Value*

I should like to conclude this chapter with some observations on the relationship between goodness, normative force and life's meaning or value. I have argued that insofar as one is engaged in a practice, one must follow the rules and that this allows one to act well in light of the goal or purpose of the practice. Only by playing chess, can one play chess well. But what do we say to the person who does not care a fig for playing chess? One response is to consider in what sense playing chess is a good in light of our further practical commitments and finally in terms of our non-contingent activities. We might argue that chess is conducive to mental health, we might say it is a pleasurable activity and that games of chess can be beautiful in something like the sense that mathematical proofs can be beautiful. In other words, we can undertake the ethical task of locating chess within the structure of a flourishing life. But, as well as wanting to know how the practice connects with our other concerns, we should like to be able to respond with an account of the distinctive value of the practice of chess. Certainly we can talk about the goods internal to the practice, but what about the goodness of the practice as a whole?

Following my discussion of Philippa Foot's account of natural goodness, I argued that we should not rest content with the thought that some trait is good in relation to our distinctive way of life. Rather, we should be able to reflectively endorse the pursuit of particular goods. I argued that this is an important part of what it is to be rational. I also mentioned that there is a sense of 'good' that is not covered by the primary attributive use identified by Geach. I had in mind the sort of usage David Wiggins describes as follows:

What the philosophical lexicographer recapitulates under "good", we might say, is the history of our constant interrogation of the life that we lead and the place where we lead it, our constant interrogation of the things that concern us or might concern us or ought to concern us. (2009: 198)

Although Geach was right that 'good' typically functions as an attributive adjective, we sometimes use the word 'good' in a rather different sense to mean the sort of thing that is worthy of pursuit all-things-considered, or the sort of thing that finally stands up to interrogation. This is what we mean when we say things such as 'pleasure is good', 'friendship is good' or 'health is good'. We do not mean that these things are good of their kind, or that they are instrumentally valuable, but that they are good *things* and worthy of pursuit. We mean, in other words, that they are final goods, standards against which actions can be measured and plausible candidates for being the end, or among the ends, of a flourishing life. The fact that a claim is being made concerning final value is indicated by the fact that 'good' functions as a predicative adjective in this context.<sup>37</sup>

So, when we consider whether to reflectively endorse a practice such as chess, we can consider whether the practice is good in the predicative sense; whether it is worth pursuing *tout court*. Because chess is a contingent practice, we can adopt an external perspective on the practice and consider its value in light of our other reasons. The more pressing case, however, is morality. As above, I mean by 'morality' the preconditions for engaging in the sorts of activities that constitute a flourishing life. Unlike chess, we cannot choose to stop engaging in

---

<sup>37</sup> Geach dismisses predicative uses of 'good' as 'a peculiarly philosophical use of words' (1956: 36), but this looks like a case of *ad hoc* monster-barring. A quick internet search shows that claims such as 'friendship is good' and 'pleasure is good' are in common parlance.

moral activity and pick up some other activity instead. We can choose to knock ourselves out with drink or drugs, but, setting aside questions of addiction, these are still actions for which we are morally accountable. Whereas we engage in the activity of chess in virtue of prior choices and contingent circumstances, we are engaged in moral activity in virtue of our existence as rational agents. As being moral agents is our plight, to ask whether morality or other forms of necessary activity are good in the no-holds-barred predicative sense can only be to ask whether life for beings like us is worthy of pursuit. In other words, it is to enquire into the human condition and whether life is meaningful.

Of course, death is a possibility; we can cease being rational agents and thus cease to be engaged in normatively governed activity. As is well known, Kant objected to suicide on the following grounds:

To annihilate the subject of morality in one's person is to root out the existence of morality itself from the world as far as one can, even though morality is an end in itself. (*Metaphysics of Morals*, 423)

I am not interested here in the case of suicide as euthanasia, but rather in suicide as the rejection of normativity, or what we might call 'existential suicide'. As Kant suggests, this is a move towards annihilating value. But, what if normativity in the world, with its various preconditions and ends, is not valuable to *me*? What if I find flourishing pointless, meaningless or absurd? Why then should I be bound by norms instead of committed to suicide? Consider Tolstoy's account of his doubts concerning the ultimate meaning or value of various activities. He recalls that at the height of his literary success, financially secure and with a loving wife and family, the question kept pressing upon him 'Why? Well, and then?' (1987: 10) whenever he considered any prospective good. Repeatedly hitting upon this question, he reported:

I felt that what I was standing on had given way, that I had no foundation to stand on, that that which I had lived by no longer existed, and that I had nothing to live by. (1987: 11)

According to the view I have been defending, rationality is a matter of recognising and responding to normative reasons. So, we can and should argue that existential suicide is irrational. But, as existential suicide is a rejection of

normative reasons quite generally, this will not carry any dialectical force. Perhaps Wittgenstein had similar thoughts in mind when he remarked:

If suicide is allowed then everything is allowed. If anything is not allowed then suicide is not allowed. This throws a light on the nature of ethics, for suicide is, so to speak, the elementary sin. (1979: 91)

Camus also wrote that ‘there is but one truly serious philosophical question and that is suicide’ (1955: 11). This may overstate the case, but it is certainly a question that arises in response to an important challenge. What then can we say to assuage doubts that the flourishing life may not be good in the gloves-off predicative sense?

It seems there are two sorts of considerations we might offer. I have attempted to account for the normative structure governing our various practices and activities, and in so doing provide an account of the good and the right. But we can still question whether any of this is worthy of pursuit because we can question whether any of it matters or is meaningful in the grander scheme of things.<sup>38</sup> Since questions of meaning are questions of significance or fit, we can proceed along two axes, which we might think of as horizontal and vertical. On the one hand, activity can be meaningful when it measures up to a higher, more encompassing standard. For instance, I explained the value of chess moves in terms of their place in a wider practice. That practice is in turn embedded in larger structures as I indicated when discussing the place of chess in a flourishing life. If we want to progress further along the vertical axis of fit, however, we must find something that both transcends and justifies human activity. Thus,

---

<sup>38</sup> It is sometimes suggested that questions about the meaning of life originate in a sense of our smallness in relation to the cosmos. As Cottingham explains the concern: ‘We humans pride ourselves on our intellectual and cultural achievements, but against the backdrop of unimaginable aeons of time through which clouds of incandescent hydrogen expand without limit, we are a strange temporary accident, no more significant than a slime or mould that forms for a few years or decades on a barren rockface and then is seen no more’ (2003: 3–4). The dizzying attempt to view our lives from an Archimedean point, or what Nagel calls the view from nowhere, leaves us with a sense of the absurd. However, it is not our smallness as such that is disturbing. If we dislike the thought of a short meaningless life confined to a tiny corner of the universe, a long meaningless life with a massive body, or in a much smaller cosmos hardly seems like a more attractive alternative. Rather, the concern seems to be that projects which loom large in our lives will appear insignificant from the point of view of the universe. However, it is highly doubtful whether we have a firm enough grasp of our lives *sub specie aeternitatis* to make sense of this claim. As Scarre observes, ‘perhaps from that standpoint [our lives] will appear not as petty and worthless but as precious jewels in a waste of material processes. Who, besides God, can say?’ (2001: 214).

Tolstoy partly assuaged his concerns, and alleviated his depression, by finding meaning in the design plan of a benevolent God.

Of course, Tolstoy's solution depends on an appeal to supernatural agency that many of us would reject. Setting this aside, however, a source of vertical meaning still may not be sufficient to address our concerns. At least, not just any source of vertical meaning would do. For example, God's existence would be little consolation if we were the equivalent of Harvard mice or broiler chickens in the Divine plan. For we are looking for a source of meaning that we can reflectively endorse. The difficulty, however, is that unless normativity is internal to human practices or activities, the reasons in favour of the practices are likely to be the wrong sort of reasons. If the reason that our practices and activities are justified is that they are conducive to some Divine plan, then the reason to engage in them just is that further purpose. As I argued in the previous chapter, if the normative reason for an action or belief is its relation to some further end, then we do not have two reasons, i.e. the end and the fact that the means produces the end. If this is right, we would be radically mistaken in thinking that we have moral reasons, prudential reasons, epistemic reasons, aesthetic reasons and reasons associated with various contingent practices. The reason to meet someone's non-contingent needs, to provide for one's future, to create and preserve the beautiful, to believe what is true, and even to exchange down to an endgame with a queenside pawn majority would finally be their contribution to a Divine plan. In other words, we would only have Divine reasons. But, then, a Divine reason is not necessarily a human reason (and if it is a human reason then it is redundant). So, it is unclear why we should reflectively endorse it. Moreover, most people most of the time do not *think* that they are acting for Divine reasons. So, the view under discussion would have the unattractive consequence that we are generally irrational in the sense that the reasons that justify our actions are rarely the reasons that actually motivate our actions.

For the above reasons, I am not convinced that we should look for life's meaning or value in something beyond the field of human practices and activities. If this is right, we are left to consider the horizontal axis of fit; the way in which our



various activities fit together *within* the field of practices and activities. For example, we find activity meaningless when it fails against the standards of our core projects, activities and practices; those to which we are most fully committed and which make us who we are. We might think of the absurdity and alienation of work which is at odds with what matters most to us or which is simply unrelated to our goals. On this view, there is nothing more to the question of the meaning of life than the question of the way in which the various parts of a life relate to one another. Judging this means evaluating practices and activities in light of each other, and, in particular, evaluating contingent practices in light of non-contingent activities such as moral requirements and the pursuit of truth. This means considering the reasons that are given by the right and the good associated with the preconditions and goals of practices and activities. It also means considering whether an agent's practical identity fits with her more idiosyncratic capacities.

It might be objected that we cannot appeal to facts about the significance of activities within a human life when it is the significance of those very facts which is under scrutiny. Indeed, it is true that we face a very general dilemma when it comes to the justification of reasons. For we either appeal to reasons of the same sort in which we case we are guilty of circular reasoning. Or, we can appeal to reasons of a different sort, such as Divine reasons, in which case it seems that the reasons will belong to the wrong category. Perhaps the most famous example of this sort of dilemma is Hume's problem of induction.<sup>39</sup> On the one hand, it is circular to justify induction on inductive grounds. On the other hand, it is inappropriate to justify induction on deductive grounds. A number of commentators have concluded that all Hume succeeded in showing is that induction is not deduction, and that it is wrongheaded to try to explain the one in terms of the other.<sup>40</sup> So, although the objection that horizontal meaning is viciously circular should be taken seriously, the alternative seems to be the misguided attempt to reduce the meaning internal to a human life to a type of meaning external to a human life.

---

<sup>39</sup> Weintraub (1995) argues persuasively that Hume's problem of induction is in fact just a special case of the ancient problem of the criterion. The paradox of analysis also has the same structure.

<sup>40</sup> See, for example, Strawson (1952: 250).

We are left with the conclusion that we can *explain* meaning within the field of practices, but that we cannot *justify* it on pain of circularity or irrelevance. But even if we cannot justify a life as a whole, perhaps it is enough to show that there is value within it. Perhaps we can do no more than identify life's sources of value and then attempt the difficult ethical work of trying to understand how to structure a life towards the end of flourishing. Finally, I see no alternative but to maintain that the meaning of life is just whatever meaning there is to be found in a flourishing life consisting of a coherent set of valuable practices and activities flowing from the distinctive activities of beings like us. There is no shortcut to answering whether life is good in the predicative sense. We must carry out the painstaking ethical work of considering what there is to be said in favour of particular practices and courses of action and how they can be integrated into a life with horizontal meaning.

## **Chapter Seven** **Truth, Virtue and Epistemic Normativity**

The previous chapter was perhaps the most ambitious of this dissertation. In response to the difficulties facing naturalistic accounts of normativity we encountered in earlier chapters, I defended the broad outlines of my own account. I considered the case of the contingent practice of chess in some detail and I touched on questions of morality and finally meaning in life. In this chapter, I defend the view that truth is the regulative norm of epistemic activity, draw out conclusions with respect to epistemic justification and respond to a number of possible objections. The emergent view stands in contrast to the instrumentalist position I discussed in chapter five and the psychologism I discussed in chapter four. It also constitutes an important test case for the account of normativity I defended in the previous chapter.

It is agreed on all hands that there is some connection between what is true and what we ought to believe. Broadly speaking for the present, it seems that we do well if we believe what is true, and we do badly if we believe what is false. However, bringing out the precise nature of this relationship is not easy. In section 7.1, I outline my view of this relationship, drawing on insights from virtue epistemology and contrasting my position with Linda Zagzebski's. Briefly, I argue that we ought to believe propositions that are produced by virtuous dispositions of thought which are directed towards and reliably conducive to truth for beings like us in an environment like ours. In sections 7.2 and 7.3 below, I consider arguments that appear to count against this view. In section 7.2, I consider an argument adapted from Bykvist and Hattiangadi (2007), which seems to suggest that my account places impossible demands upon epistemic agents. I argue that this is not the case, drawing on the distinction between what we ought to believe and what it would be good for us to believe. In section 7.3, I look at an argument suggestive of the idea that epistemic normativity depends in some measure upon practical interests. In particular, I consider a thought experiment of Jason Stanley's (2005) which intuitively suggests that ascriptions of knowledge depend upon practical interests. I accept that our intuitions tend in that direction, but attempt to explain away the experiences by disentangling competing sources of normativity.

### 7.1 True Beliefs and Epistemic Virtue

Part of the view I would like to defend is that the activity of belief-formation is necessarily governed by the norm of truth. For epistemic agents it is constitutive of having beliefs that they are governed by epistemic norms, and in particular the epistemic goal of truth. It is not that agents form beliefs with the aim of believing true propositions, as one might have the aim of believing, say, interesting propositions.<sup>1</sup> Rather, part of what it is to be a belief is to be normatively governed by truth. Just as winning is the goal of chess, narrowly construed, and flourishing is the goal of human action quite generally, truth is the goal of belief-formation. Like chess, belief-formation can be evaluated in the light of other standards. Beliefs, or rather their content, can be beautiful, useful and the like. But, like morality, it is the inescapable plight of beings like us to form beliefs. Humans are essentially epistemic agents and there could not be an epistemic agent whose beliefs were not governed by and evaluable against this normative standard. Indeed, it would be entirely opaque to call a mental state a ‘belief’ if it were not primarily evaluable against the standard of representing the world as it is.

In the last chapter I discussed Moore’s paradox in the context of goodness in chess. Something has gone amiss with an agent who sincerely asserts ‘*p*, but I don’t believe that *p*’. This is despite the fact that the assertion is not only logically consistent, but is true in many circumstances. Given human fallibility, it is close to certain that some of an individual’s beliefs are false even though they are not in a position to identify which ones with certainty. At least it is necessarily true that if we take two epistemic agents with contradictory beliefs, at least one of them holds a false belief. As I suggested in the previous chapter, at

---

<sup>1</sup> It is often said that belief aims at truth, and this can be a helpful metaphor as long as one does not confuse it with the proposition that epistemic agents necessarily aim at truth. For instance, Steglich-Peterson (2009: 396) writes: ‘It is a conceptual truth about believing, that someone believing *p* aims at believing *p* truly, it is simply impossible to consciously and deliberately believe that *p* despite evidence to the contrary.’ But this is not a conceptual truth and it is probably not even a truth. For even if, as Steglich-Peterson suggests, it is impossible consciously and deliberately to believe that *p* despite evidence to the contrary, it does not follow that believing truly is among the aims of an agent who believes that *p*. It seems perfectly possible to believe *p* despite one’s aims. I have argued that there is always a reason to believe a true proposition, but it would be rather optimistic to think that we always aim at what we have reason to do.

least part of the problem is that the assertion of a Moorean sentence must be understood as a violation of the norm of truth that governs belief. The assertion implicitly denies that in holding beliefs one is normatively committed to the goal of holding true beliefs. We can make no sense of this denial, hence the assertion's paradoxical appearance.<sup>2</sup>

The fact that the activity of belief-formation is governed by the norm of truth also helps to explain why from the first-person deliberative standpoint the question of how I believe the world to stand collapses into the question of how the world in fact stands.<sup>3</sup> As Gareth Evans noted:

In making a self-ascription of belief, one's eyes are, so to speak, or occasionally literally, directed outward – upon the world. If someone asks me 'Do you think there is going to be a third world war?', I must attend, in answering him, to precisely the same outward phenomena as I would attend to if I were answering the question 'Will there be a third world war?'. (1982: 225).

This phenomenon has come to be known as transparency.<sup>4</sup> Other propositional attitudes are not similarly transparent. We do not usually direct our eyes outward in order to determine whether we desire that *p*. This is because we do not need evidence for *p* in order to be rationally entitled to desire that *p*; normally the reverse is true. By contrast, one is only rationally entitled to believe that *p* if one is rationally entitled to take *p* to be true. So, on the assumption of one's own rationality, one self-ascribes a belief that *p* by considering the evidence that *p* is true. Of course, when it comes to beliefs we have already formed, the assumption of one's own rationality may be unwarranted. Nevertheless, when we come to self-ascribe a belief, we assume that it was formed on the basis of the evidence for its truth, for we cannot make sense of ourselves unless we assume that we are rational.

---

<sup>2</sup> This has certain formal affinities with transcendental arguments. So, the thought runs, it is manifestly true that we understand and engage with rational others, but for this to be possible it must be the case that their beliefs are normatively regulated by the goal of truth.

<sup>3</sup> Similarly, the question 'What do I want?' seems to collapse into the question of 'What is all things considered good?' See Moran (1988: 145) for further discussion.

<sup>4</sup> See, for example, Shah (2003: 448). Although Evans' idea is widely accepted, see, Cassam (2011) for several concerns about the transparency thesis.

Another consideration pointing in the same direction is that the norms governing rational inferences between beliefs are the same as the rules of deductive logic which guarantee truth preservation. So, for example, if *S* believes *p* and *S* believes *q*, then *S* is rationally committed to believing *p* and *q*. If *S* believes that if *p* then *q* and *S* believes that *p* then *S* is rationally committed to believing that *q*.<sup>5</sup> If, however, belief were not normatively governed or regulated by truth, then it is obscure why agents should be rationally committed to these norms. Propositional attitude that are not governed by the norm of truth do not behave in this way. For instance, my desire that *p* and my desire that *q* do not jointly commit me to a desire that *p* and *q*. I might be caught between my simultaneous desires for tea and coffee, without thereby desiring both. Similarly, if I entertain the thought that if *p* then *q* and I entertain the thought that *p*, I am not rationally committed to entertaining the thought that *q*. Perhaps *q* is too terrible to contemplate.

Once we grasp what is meant by sincerely asserting a belief that *p*, we cannot without conceptual confusion wonder whether the speaker also believes that *p* is true. To employ another of Moore's ideas, the questions 'I believe that *p*, but is *p* true?' and '*S* (sincerely) asserts that *p*, but does *S* believe that *p* is true' are closed. To understand what it is to believe that *p* is to understand that whoever sincerely asserts *p* believes *p* and whoever believes *p* believes *p* to be true. This contrasts with open questions such as '*p* is good / beautiful / promotes my self-interest, but do I believe that *p*?' Our handle on the concept of belief depends on the relationship in which belief stands to the goal of truth, not the contingent relations which beliefs bear to other valuable goals. The beauty or goodness of a proposition can certainly provide us with a reason to believe it, but these are aesthetic or ethical reasons. From a purely epistemic perspective, any such reason is always over-ridden by the proposition's falsity.

---

<sup>5</sup> Of course, the fact that I am rationally committed to the conclusion does not mean that I will in fact believe the conclusion. For, as Carroll (1895) illustrated with his example of Achilles and the Tortoise, I can always question the normative force of *modus ponens* or any other rule of logical inference. Normative force is not motivational force and Logic does not take one by the neck and force one to believe as Achilles hoped.

Although it is not possible to find a rational interpretation for Moorean sentences like ‘ $p$ , but I do not believe that  $p$ ’, we can make sense of the assertion ‘ $p$ , but I desire not to believe that  $p$ ’. A wife might desire to believe that her philandering husband is faithful. Caught in this position, she might even say ‘my husband is unfaithful, but I don’t believe it’, and we would have a sympathetic grasp of her meaning. In fact, she does believe that her husband is unfaithful, for otherwise we could make no sense of her sincere assertion to that effect, but she desires, and perhaps intends, to form the contrary belief. Her utterance may well be part of that very strategy. It is familiar enough that repetition of a falsehood can lead to belief in its truth. The most convincing liars, it is sometimes said, finally come to believe their own lies. In order to fulfil her desire to falsely believe in her husband’s fidelity she may also avoid looking for compelling evidence. Or she might actively seek scraps of evidence to the contrary. However, in doing so she is deliberately cultivating an epistemic vice. This does not entail that her behaviour is wrong all things considered. Moral, prudential, aesthetic and other sorts of reasons can trump epistemic reasons on occasion. Nevertheless, she is manifesting an epistemic vice.

Claiming that the truth-goal is always relevant to epistemic appraisal is neither to say that having a true belief is necessary, nor sufficient, for being the appropriate object of epistemic praise. The reason for this is that we do not *directly* evaluate beliefs against the normative standard of truth. This is clear from the fact that we neither blame someone who is subject to epistemic misfortune, nor praise someone who is the beneficiary of epistemic good luck. The world can be a cruel and deceptive place. Zebras can turn out to be cleverly painted mules. Epistemic luck can trump diligence. So it cannot be quite right to say, as I suggested as a first pass, that we always do well if we hold true beliefs, or that we always do badly if we hold false beliefs.

Moreover, we do not have immediate or direct control over our beliefs. If I see a red truck coming towards me I cannot by an act of will believe that it is a yellow tractor. Generally, if I am faced with misleading but compelling evidence, I

cannot simply choose to disregard it and believe the truth.<sup>6</sup> But although we cannot always make an immediate decision to believe or withhold belief from some proposition, we can aim to bring about the appropriate circumstances for the correct beliefs to arise.<sup>7</sup> We can do this in at least two ways, by working on ourselves directly or indirectly.<sup>8</sup> On the one hand, we can affect our mechanisms of belief formation by cultivating epistemic virtues and improving our background knowledge. On the other hand, we can influence the input. This can mean changing our location, focusing on different aspects of the world or seeking out different sources of testimony. It can also mean actively exploring and engaging with our environment.<sup>9</sup> We do this when we shuffle tiles on a Scrabble rack, complete a jigsaw puzzle and try to set a philosophical argument down on paper.

Chisholm has suggested that the situation of the epistemic agent with respect to his beliefs is analogous to a debtor who has an obligation to repay borrowed money, but not at the moment at which he takes on the obligation (1991: 126). Chisholm sees this primarily as a matter of discharging our epistemic obligations in due course, by carefully reflecting upon our existing stock of beliefs. However, to stretch the analogy a little, it seems more important that the debtor ensures that he is in a position of fiscal responsibility before he takes on the loan, so as to avoid taking on an obligation he cannot discharge. To this end, I suggest that a virtuous epistemic agent should cultivate those dispositions of thought

---

<sup>6</sup> As William Alston argues, ‘If I were to set out to bring myself into a state of belief that *p*, just by an act of will, I might assert that *p* with an expression of conviction, or dwell favourably on the idea that *p*, or imagine a sentence expressing *p* emblazoned in the heavens with an angelic chorus in the background intoning the Kyrie of Mozart’s Coronation Mass. All this I can do at will, but none of this amounts to taking on the belief that *p*’ (1988: 263). There is of course a large literature on the subject of doxastic voluntarism. For a sampling, see Naylor (1985), Feldman (2000), Katzoff (2001), Booth (2007) and Hieronymi (2008). Even if it is true that we cannot form beliefs at will, there are interesting questions concerning what sort of impossibility is involved. Is it ‘mere’ psychological impossibility, or is it conceptual, metaphysical or logical impossibility? For more on this see Bennett (1990).

<sup>7</sup> Huss (2009: 257–61) lists ten ways in which we can exert indirect control over our beliefs, but they are variations on the two themes I mention. See also Nottelmann (2008: 84–93) for further detailed discussion of doxastic control.

<sup>8</sup> A third possibility is to bring it about that we believe that *p* by bringing it about that *p*. However, this has little to do with epistemic normativity, which is a matter of aligning one’s beliefs to the world as it is.

<sup>9</sup> As Noë during his excellent discussion of the role of action in belief-formation, ‘the process of perceiving, of finding out how things are, is a process of meeting the world; it is an activity of skilful exploration’ (2004: 165).



which are reliably conducive to the acquisition of true beliefs. These are epistemic virtues such as intellectual integrity, precision, care and consistency; virtues because of the relationship in which they stand to truth. By contrast a vicious epistemic agent either fails to cultivate epistemic virtues, or deliberately adopts strategies which are antithetical to truth. Determining the particular dispositions of thought which are epistemically virtuous or vicious is formally a matter for empirical psychology, for it is a matter of determining which dispositions of thought are reliably conducive to forming true beliefs for beings like us in an environment such as our own. However, in practical terms, the common sense of responsible and mature epistemic agents is normally a sufficient guide.

So the account of epistemic evaluation I favour is virtue-based. When we transfer our attention to the prior cultivation of epistemic virtue, we can see that the fact that we cannot decide to believe at will is not a threat to the idea that we are epistemically responsible for our beliefs. Consider again the example of the red truck coming towards us. If we automatically form the belief that there is a red truck, then we are in fact manifesting an epistemic virtue, namely the disposition to believe the evidence of our senses in the absence of obvious defeaters. If we lived in a world in which apparently red trucks routinely turned out to be yellow tractors on closer inspection, then it is quite plausible to suppose that we would not be disposed to form the belief that there is a red truck coming towards us. We do not, for example, form beliefs about tiny trucks in the distance, although that is arguably what we see. Epistemic virtues are stable dispositions to form certain sorts of belief in certain circumstances. It is no concern that we form beliefs automatically and unreflectively. Doing so requires the sort of skill that is the manifestation of epistemic virtue.

So, a virtue-based account of epistemic evaluation is attractive because it enables us to account for doxastic responsibility. It also issues in intuitively correct judgments in cases of epistemic luck. We do not do badly, epistemically speaking, if we exercise a virtue of thought that unluckily produces a false belief and the same is true, *mutatis mutandis*, for epistemic vice. Moreover, a virtue-based account helps to resolve a third issue. If truth were the end of belief-

formation, it is sometimes objected, one way to satisfy this end would be simply to form a great many beliefs. Thus, van Fraassen argues, we would do well to believe everything we can believe to ‘be sure to catch the truths in our net’ (2002: 86). However, this is a terrible epistemic strategy. By casting our cognitive nets as wide as possible we would also catch a great many falsehoods. So, the argument proceeds, we must have at least two cognitive aims, namely the acquisition of true beliefs *and* the avoidance of false beliefs. However, as van Fraassen argues, the aims of holding true beliefs and avoiding false beliefs are distinct and involve trade-offs; ‘they cannot be jointly maximised; to some extent, each is gotten at the expense of the other’ (2002: 87). For, he argues, if avoidance of error were our primary aim, we would do well to pare down our beliefs to a minimum of secure beliefs; perhaps, he suggests, tautologies. This suggests some sort of compromise between the acquisition of true beliefs and the avoidance of false beliefs, rather than truth as an overriding aim.

A virtue-based account has a simple response to this line of argument, namely that believing as many propositions as possible is not the manifestation of an epistemic virtue. One straightforward reason is that believing as many propositions as possible is not reliably conducive to truth. But this is not quite to get to the heart of the matter. Let us say that a stable disposition of character counts as the virtue of justice if it is reliably conducive to giving people what they deserve or are owed. Imagine, then, that someone claimed that justice was a matter of wandering around beating everyone one encounters with a stick. This might be defended on the grounds that people are generally wicked and deserve a beating. Even if this is so, and people were generally getting what they deserved at the hands of the stick-wielding vigilante, we should hardly count this as administering justice. The reason for this is that actions are virtuous when they are skilfully directed towards their targets. Similarly, if one forms as many beliefs as possible one would likely catch many true propositions in one’s cognitive net. But one would not believe them *because* they were true and so one would not be acting on the basis of epistemic virtue.

For the above reasons, I think virtues are the primary focus of epistemic evaluation. But although I am drawing on ideas from the field of virtue

epistemology, I am not advocating what Zagzebski describes as a ‘pure virtue theory’. So, Zagzebski describes pure virtue theories as follows:

The concept of a right act is *defined* in terms of the concept of a virtue or a component of virtue such as motivations. Furthermore, the *property* of rightness is something that emerges from the inner states of persons. (1996: 79)

By contrast, I am defending the view that an agent is epistemically virtuous because his stable dispositions of thought are reliably conducive to the good, i.e. truth or cognitive contact with reality. A stable disposition of character is a virtue because of the relation in which it stands to the epistemic good of true beliefs, which is the final source of normativity. True beliefs are not good because they are the product of virtuous activity.

One reason to prefer this order of explanation is the difficulty involved in understanding what counts as an epistemic virtue without a prior understanding of the epistemic good. Zagzebski claims that ‘an act is right because it is an act a virtuous person would (or might) do’ (1996: 80). However, without a conception of the good, it is far from clear how we could recognise a virtuous person. Aristotle advised us to become wise by observing the wise. However, it is often asked, how are we to identify the wise (*phronimoi*) without possessing practical wisdom (*phronesis*) in the first place? Although this is a very old, well-known problem, it is difficult to see any satisfactory solution. Zagzebski appeals to the example of ‘persons whose goodness shines forth from the depths of their being’ (1996: 83), but it is hard to know quite what to make of this. She fleshes out the comment as follows:

I believe it is possible that we can see the goodness of a person in [a] rather direct way. She may simply exude a “glow” of nobility or fineness of character, or as I have occasionally seen in a longtime member of a contemplative religious order, there may be an inner peace that can be perceived to be good directly, not simply because it can be explained on the theoretical level as a component of *eudaimonia*. (1996: 83)

I do not doubt that we occasionally have the experience of such encounters. Nor do I doubt that the actions of virtuous others are sometimes our best guide to the virtuous course of action. This is especially the case during our early moral and

epistemic development. However, I do not see how Zagzebski's comments could help us to recognise a virtuous person without some understanding of how their behaviour relates to the good. Although we might be subject to the kind of impression Zagzebski describes, appearances can be deceptive. The virtuous do not glow, and sadly they may appear rather dull. By contrast, the vicious can exude inner peace. The devil, as they say, has all the best tunes. We sometimes communicate our dispositions through our demeanour, but this is a highly unreliable guide to virtue. Much better, I contend, to evaluate a person's dispositions in the light of the good.

### *7.2 Impossible Standards: obligation and evaluation*

I have argued, then, that an epistemic agent is always answerable to the reliability of her cognitive processes with respect to forming true beliefs. I would also like to endorse the stronger claim that this is the *only* criterion against which it is appropriate to evaluate epistemic agents *qua* epistemic agents. In order to develop this idea, it will be helpful to consider a near alternative and why it fails.

Bykvist and Hattiangadi criticise the following putative principle:

For any  $S, p$ :  $S$  ought to believe that  $p$  if and only if  $p$  is true. (2007: 277)

So, according to this principle, the psychological state of believing some proposition is correct, fitting or appropriate if and only if the proposition is true. If a proposition is false, one ought not to believe it regardless of other considerations in its favour.<sup>10</sup> Unfortunately, as Bykvist and Hattiangadi observe, this principle quickly runs into serious trouble. For, there are clearly true propositions for which it is not the case that one ought to believe them. For example, there are true propositions which one *cannot* believe. No human mind can entertain a proposition containing infinitely many true conjuncts, and, hence,

---

<sup>10</sup> Again, this is not to suggest that believing the truth is always what we ought to do in an all-things-considered sense. As I conceded in chapter five, situations occur in which it is best to believe what is false. We may have moral, aesthetic or prudential reasons to believe what is untrue, and these reasons may trump our epistemic reasons. So we need to read the 'ought' as purely epistemic.

no human can believe such a proposition despite its truth. As ‘ought’ implies ‘can’, no human ought to believe such a proposition.<sup>11</sup>

Now, the view I am advocating does not face quite the same difficulty because, on my view, it is not true that we always ought to believe what is true. Imagine that I see what appears to be a red surface in ordinary lighting conditions, and so form the belief ‘there is a red surface over there’. But, in fact, what I see is due to a hidden jokester projecting red light on to a white surface.<sup>12</sup> Given this outré scenario, my view would entail that it is epistemically virtuous to believe what is false. This is because trusting one’s senses, in the absence of any reason to suppose that there is a defeater for one’s perceptual evidence, such as a jokester, in the vicinity, is reliably conducive to forming true beliefs for beings like us in an environment such as ours. We would have, as always, a reason to believe the truth, but that reason would be cognitively unavailable for non-culpable reasons. This seems the correct response insofar as the victim of the joke is not intuitively blameworthy when he forms the false belief that he sees a red surface, but yet something *has* gone wrong with regard to his belief.

Nevertheless, it is not difficult to envisage an analogous objection. Given that epistemic virtue is understood in terms of reliability with respect to true beliefs, ought not an agent cultivate the disposition of epistemic infallibility? Ought not a virtuous epistemic agent believe the truth and nothing but the truth? But, for the reasons noted above, infallibility is impossible for beings like us. We cannot believe every true proposition, because we cannot even entertain or consider every true proposition. So, doesn’t my position also saddle epistemic agents with obligations they are impotent to discharge?

I think not, for reasons I’ll now explain. In this chapter, I have been referring more or less interchangeably to epistemic evaluation and what an epistemic agent ought to do. But, as we saw in the last chapter, the good and the right are

---

<sup>11</sup> Whiting argues that Bykvist’s and Hattiangadi’s putative principle is too strong and that from the fact that the correctness of a belief depends on its truth we should draw the conclusion that ‘For any *S*, *p*: *S* may believe that *p* if and only if it is true that *p*’ (2010: 216). But although I agree with Whiting that the principle involving ‘ought’ is too strong, I think that Whiting’s principle involving ‘can’ is too weak. I argue for middle ground below.

<sup>12</sup> I borrow this example from Ernest Sosa (2007: 31–4).

distinct. Epistemic obligation or rightness, what one ought to do, is associated with the necessary preconditions for epistemic agency. These include specifically epistemic conditions such as conformity to the basic principles of logic and also conditions such as free will, which are necessary for rational agency quite generally. These are not, however, standards against which we can evaluate epistemic agents, but preconditions for evaluation. We cannot criticise non-rational representational systems, such as a computers and all or most nonhuman animals, for failing to form true beliefs. We can only evaluate the goodness of artefacts and all or most nonhuman animals, against their function and life-form respectively. We cannot evaluate them epistemically because 'ought' implies 'can' and non-rational cognitive systems cannot believe other than they do. Even saying this much assumes that we can coherently ascribe beliefs to non-rational beings, which is questionable. So, for a creature to meet the necessary conditions of epistemic agency is for it to be the sort of creature which is appropriately subject to epistemic evaluation. Epistemic evaluation then focuses on the belief-forming mechanisms of the agent. Insofar as the agent's belief-forming mechanisms are reliably conducive to forming true beliefs, the agent is epistemically virtuous, or, in an alternative idiom, he does well epistemically.

The answer, then, is that it would be good, excellent, indeed ideal, to cultivate the disposition of infallibility. This is why omniscience has traditionally been attributed to the God of the Judeo-Christian-Islamic tradition. However, we are under no *obligation* to do so. This does mean, it is true, that insofar as we are unable to believe every truth we are to some extent epistemically bad or imperfect. Yet this much is commonplace. It is perfectly obvious that we are all fallible and, to that extent, less than ideal. This also coheres with my argument in chapter five that there is always *some* epistemic reason to believe trivial true propositions. If we are inclined to deny that a person should be evaluated negatively due to the unavoidable cognitive limitations, then it is because of the connection between evaluation and blame. To evaluate someone as bad in some respect is frequently not only to label them imperfect, but also to censure or to reprimand. But censure is inappropriate when someone fails to do what is best because they cannot do so. Although blame can rightfully accompany failures to

be good as well as failures to discharge obligations, in both cases it is fitting only for someone who could have done otherwise. For someone who could not have done otherwise, it is appropriate to merely withhold admiration.

### *7.3 Knowledge and Practical Interests*

If the above arguments are compelling, the position outlined in section one emerges in tact. However, as we shall see shortly, there is some reason to suppose that epistemic evaluation depends not only on a belief's (indirect) relationship to truth, but also on practical interests. At least, whether or not we are intuitively inclined to *ascribe* knowledge seems to depend upon relevant practical interests. This is of course related to the view that we looked at in chapter five that epistemic normativity is instrumental normativity. Nevertheless, as is so often the case in philosophy, we can improve our understanding of a question by addressing it from several perspectives. Moreover, we can now draw on insights from the positive account of epistemic normativity developed over the past two chapters.

Given my concern with the nature of epistemic normativity, it might seem surprising that I have said so little about the nature of knowledge. However, several of the elements of my view are in place. Since Plato it has been widely thought that knowledge is true belief plus some normative element such as warrant or justification. There must be some normative element involved because knowledge is valuable, and subject to epistemic evaluation, in a way that mere true belief is not. Now, I have argued that the evaluative element of knowledge depends upon epistemic virtue and its relationship to truth. So, roughly, on my account, I know some proposition  $p$  if  $p$  is true, if I believe that  $p$  and if I have acquired  $p$  virtuously. This is not an analysis, but a list of three necessary conditions. Nevertheless, I think it accounts for most of what we want to say about knowledge.<sup>13</sup> However, the point for present purposes is that I take the difference between knowledge and mere true belief to be something normative associated with truth as the end of epistemic activity.

---

<sup>13</sup> For a far more detailed account of knowledge in the virtue-theoretic tradition, and a proposal for how to deal with Gettier cases, see Zagzebski (1996: 293–9).

Let's consider a reason to suppose that I am mistaken about this. Jason Stanley describes the following two scenarios:

*Low Stakes*: Hannah and her wife Sarah are driving home on a Friday afternoon. They plan to stop at the bank on the way home to deposit their paychecks. It is not important to do so, as they have no impending bills. But as they drive past their bank, they notice that the lines inside are very long, as they often are on Friday afternoons. Hannah says, 'I know the bank will open tomorrow, since I was there just two weeks ago on Saturday morning. So we can deposit our paychecks tomorrow morning.'

*High Stakes*: Hannah and her wife Sarah are driving home on a Friday afternoon. They plan to stop at the bank on the way home to deposit their paychecks. Since they have an impending bill coming due, and very little in their account, it is very important that they deposit their paychecks by Saturday. Hannah notes that she was at the bank two weeks before on a Saturday morning, and it was open. But, as Sarah points out, banks do change their hours. Hannah says, 'I guess you're right, I don't know that the bank will be open tomorrow.' (2005: 3–4)

Now, Stanley suggests that Hannah is right in both cases. In the first scenario she knows that the bank will be open on Saturday morning. In the second scenario she doesn't know this. However, the only difference between the two cases is that in *High Stakes* it is important to deposit the cheque, whereas in *Low Stakes* it is not. So, if these linguistic intuitions represent reality, ascriptions of knowledge would seem to depend on practical interests.

In truth, my intuitions concerning these cases don't carry much conviction. I doubt one would feel much inclined to challenge Hannah in either case, but it seems to me that one could do so without manifesting any linguistic incompetence or conceptual confusion. Consider a variant combining the two cases:

*Low Stakes\**: Hannah and her wife Sarah are driving home on a Friday afternoon. They plan to stop at the bank on the way home to deposit their paychecks. It is not important to do so, as they have no impending bills. But as they drive past their bank, they notice that the lines inside are very long, as they often are on Friday afternoons. Hannah says, 'I know the bank will open tomorrow, since I was there just two weeks ago on Saturday morning. So we can deposit our paychecks tomorrow morning.' But, as Sarah points



out, banks do change their hours. Hannah says, 'I guess you're right, I don't know that the bank will be open tomorrow.'

Quite likely the emphasis in Hannah's last sentence would be on 'know' - 'I guess you're right, I don't *know* that the bank will be open tomorrow'. She might also accompany her assertion with a roll of the eyes and cast her mind back to philosophy classes in which she had to admit that, no, she didn't *know* that she wasn't a brain in a vat. I don't wish to deny that we are *more* inclined to say Hannah has knowledge in *Low Stakes* than in the *High Stakes*. Rather, the point is that we don't have a particularly firm intuitive grasp of the level or type of justification required for knowledge ascriptions. As indicated by *Low Stakes\**, the criteria for accepting knowledge ascriptions are not a straightforward function of the practical interests relevant to the situation. It would be a mistake, therefore, to rest substantive conclusions upon our rather vague intuitions concerning such cases. Still, I do accept that our inclination to ascribe knowledge or withhold ascriptions of knowledge on the basis of the practical interests presents a *prima facie* objection to the position I advocate.

My view, then, is that Hannah knows that the bank is open in both cases regardless of her practical interests. Knowledge does not depend upon financial solvency. Our overall evaluation of epistemic agents can depend on practical interests, but not our epistemic evaluation of epistemic agents. We need to note here that epistemic agents and their epistemic activities are subject to evaluation on a number of grounds which one can easily fail to distinguish. We are never solely engaged in epistemic pursuit. As well as being epistemic agents, we are also, for example, moral agents with certain roles in our societies. Importantly epistemic activity provides the basis for action. Instrumentalists are right that beliefs are the premises for practical syllogisms, and that true beliefs often enable us to satisfy our desires and successfully engage in valuable activities. When these activities are of significant value, we have a non-epistemic duty to form beliefs with especial diligence. W. K. Clifford (1877) was right that there are ethical reasons for belief. So, a lighthouse keeper, for example, has a duty to form true beliefs with respect to the weather. She has a duty to know when a storm is brewing. Yet, this duty is a moral duty with respect to her epistemic

behaviour, not an epistemic duty. The source of her duty is the moral obligation she has to protect the lives of seafarers.

The varieties of justification can easily become obscured when we come to ascribe knowledge to a person. So, given that a person has a true belief, we ask ourselves whether she is justified in holding that true belief before ascribing her knowledge. But the sources of justification, the sources of normativity, are diverse. One can be justified in one respect, but not another. Epistemic justification, moral justification, prudential justification and aesthetic justification, as well as the many sources of justification associated with particular contingent practices, can cut across each other in all directions. One's belief can be epistemically justified insofar as it is the product of a virtuous disposition of thought. However, the evidence for the belief might be insufficient for one to be pragmatically justified. Thus, we might be inclined to say that a glance is sufficient to know that I have tied a knot securely if I am tying a dog to a lamppost, but not if I am a surgeon tying an artery. In fact, from an epistemic perspective – the appropriate perspective from which to judge knowledge claims – one has an equal claim to knowledge in both cases (assuming that a glance is an equally reliable method). Yet, it *would* be reasonable to call into question the surgeon's belief, and we might do that by asking, "do you really *know* that the patient's artery is tied properly?" The surgeon is not pragmatically justified and he is not justified in holding his belief in an all-things-considered sense.

These examples show that ascriptions of knowledge have a useful social function.<sup>14</sup> We frequently use 'knowledge' as an honorific term to indicate that a belief is justified all-things-considered. This is a perfectly legitimate use outside the context of philosophical epistemology. Moreover, we use the term 'knowledge' as an honorific for epistemic agents. To say that someone has knowledge, that he is knowledgeable, is to praise him in a certain way. It suggests that when he acts on the basis of his beliefs we can reasonably expect him to succeed in his various pursuits. Moreover, when others act on the basis of

---

<sup>14</sup> Williams (2002: Ch. 3) develops this idea well, focusing particularly on the importance of accuracy and sincerity as social epistemic values.

his testimony they also can reasonably expect to succeed in their endeavours. As it sometimes put, knowledge is a collective good. So, to ascribe someone knowledge is to evaluate that person as a valuable member of a community. *Pace* Descartes, knowledge does not require certainty but rather a certain degree of probability or reliability based on warrant. Stating the precise degree of reliability necessary for knowledge is of course a difficult matter. But fortunately there is no need to do so for present purposes. The point is that the degree of certainty can sometimes fall below the level which we would require a reliable or trustworthy member of the community to possess for beliefs which are the basis for important purposes. This being the case, we are inclined to withhold the honorific 'knowledge' in cases where someone is sufficiently justified to know a proposition, but is insufficiently justified to act as a reliable informant for the pursuit of our practical interests. I submit that this is what explains our intuitions concerning knowledge ascriptions in the cases of *Low Stakes* and *High Stakes*.

In conclusion, this chapter has been a case study for the approach to normativity I defended in chapter six and also an alternative to the approaches to epistemic normativity I rejected in chapters four and five, namely psychologism and instrumentalism. I have taken the claim I made in the previous chapter, that the goal of epistemic practice is the formation of true beliefs and argued for it in greater detail. In particular, I have considered the implications of this view for our practices of epistemic evaluation. Thus, I have argued that the foci of epistemic evaluation are virtuous dispositions of thought and defended this view against two possible challenges. The challenge of impossible obligations allowed me to apply the distinction between the good and the right in a definite context. The challenge of practical interests allowed me to explain in more detail the way in which belief-formation can be evaluated against competing standards of normativity.

## **Chapter Eight** **Normative Force, Reasons and Desires**

I began this dissertation by proposing that we can profitably understand three influential philosophical arguments in terms of the difficulty naturalists face in locating the truth conditions for normative propositions. I suggested that the particular problem was to identify value-makers, or truth-makers with normative force. Value-makers not only justify actions and beliefs, but they also bind us and move us. This can seem mysterious and it is especially hard to explain from a naturalistic perspective. We have already seen some reasons for this and we shall explore the subject further in the final two chapters. In this chapter, however, I want to focus primarily on whether it is possible to resolve the problem of normative force by grounding reasons in desire. In fact, we shall see that a number of philosophers have thought that desires are the only possible explanation for the phenomenon of normative force since we need desires in order to explain motivation. This would be very bad news indeed for the sort of realist, externalist account of reasons I have developed in the previous two chapters. For the good-makers and right-makers I have identified are contingently related to desires at best. Fortunately for my account, there are good reasons to think that a desire-based theory of reasons does not explain normative force, as I shall argue in this chapter.

### *8.1 Desires-Based Reasons*

When I argued against instrumentalism with respect to epistemic norms in chapter five, I focussed mainly on the role of true beliefs. But, although instrumentalists typically argue that truth is the proximate goal of belief formation, they also claim to have identified a final source of normativity in the satisfaction of an agent's desires. For instance, Kornblith writes that epistemic norms are 'derived from our desires in a way that removes any mystery surrounding them' (2002: 157), and again that 'epistemic evaluation finds its natural ground in our desires in a way that makes truth something we should care about' (2002: 161). Similarly, Papineau claims:

There is always a species of derived personal value to having true beliefs that are relevant to action, for such truth will always help you to find a way of satisfying whatever desires you have. (2003a: 17)

When it comes to the final source or ground of normativity, naturalists avoid queer notions of objective, *sui generis*, final epistemic value. Instead, they turn their reflection into their own breasts and discover desires. I have already questioned whether the satisfaction of desires on the basis of true beliefs can explain the specifically epistemic dimension of epistemic normativity. In this chapter, I would like to consider whether, in any case, desires are suitable candidates to ground normativity. In this section, I look at the relationship between desires and motivation, the explanation of which is sometimes thought to require a desire-based theory of reasons (henceforth, DBR). In the next section, I look at the relationship between desires and rationality. I conclude that the two most common lines of arguments in favour of grounding normativity in desire do not work, and that DBR is inconsistent with our practices of reasoning about ends.

The question of whether reasons can depend on, or be grounded in, desires is a major fault line in metaethics. Philosophers in the ‘yes’ camp include Foot (1972), Williams (1981), Smith (1994) and Schroeder (2007). The ‘no’ camp includes Nagel (1970), Scanlon (1998), Dancy (2000), Foot (2001) and Parfit (2011). As I remarked in chapter four, I favour metaphysical externalism with respect to reasons. I take practical reasons to stand objectively in favour of actions and theoretical reasons to stand objectively in favour of beliefs. Subject to minor qualifications discussed below, I do not take it that desires stand objectively in favour of a belief or an action. Our desires produce contingent needs, such as the need to have £30 in order to watch a football match. However, it is my view that the normative force of contingent needs depends entirely on whether there are independent reasons for pursuing the end. So, I am in the ‘no’ camp.

I am not, of course, the only philosopher to be wary of grounding reasons in subjective mental states. Here is Thomas Nagel on the subject of ethical reasons:

It will not in any case do to rest the motivational influence of ethical considerations on fortuitous or escapable inclinations. Their hold on us must be deep, and it must be essentially tied to the ethical principles themselves, and to the conditions of their truth. The alternative is to

abandon the objectivity of ethics. This is a course which cannot be excluded in advance, but it should not be taken before serious attempts to rescue the subject have failed. (1970: 6)

Nagel is suggesting that grounding ethical reasons in ‘escapable inclinations’ is a method of last resort because it means abandoning objectivity. If Nagel is right, and if there is a presumption in favour of moral objectivity, then there may be a burden of proof argument against desire-based theories of reasons in the vicinity.<sup>1</sup> Moreover, it is very plausible that any presumption in favour of objectivity with respect to epistemic reasons is at least as strong as the presumption in favour of moral objectivity. Most hard-nosed naturalists would agree that the presumption is stronger.

If one is not persuaded of the merits of burden of proof arguments, or of this one in particular, it remains the case that for many philosophers the strongest argument in favour of DBR is the alleged impossibility of the alternative, namely an objective theory of reasons.<sup>2</sup> Thus, Philippa Foot, who abandoned her commitment to DBR, remarks:

I now wonder why, given the obvious indigestibility of the idea that morality is indeed a system of hypothetical imperatives, I should have accepted it even for a short time. What seemed to force it on me was the sheer difficulty of showing a practical rationality that was independent of desire or interest. (2001: 60–1)

Explaining the normative force of reasons independently of desire is indeed a difficult task. However, as we shall see, it is especially difficult for naturalists who need to identify scientifically available facts that can simultaneously explain and justify our actions.

In the remainder of this section, I shall explain why it is sometimes thought an appeal to desires can help to solve questions concerning the normative force of reasons and then offer a number of reasons to doubt that desires can do the work

---

<sup>1</sup> Burden of proof arguments are fairly common in the debate over moral realism (see, for example, Brink (1989: 23–4), Nagel (1986: 143) and Shafer-Landau (2003: Ch. 1). The debate concerning moral realism is closely related to questions of objectivity insofar as moral realists claim, among other things, that the truth conditions for normative propositions are mind-independent.

<sup>2</sup> Finlay (2007: 6) makes the same point.

required of them. In effect, this is a *tu quoque* argument designed to show that DBR has no particular advantage over an objective theory of reasons. This undermines the motivation for DBR and, in doing so, strengthens any presumptive case for an objective theory of reasons. One limitation of a *tu quoque* argument is that it will have little dialectical force for someone who does not recognise the presumptive case in favour of an objective theory of reasons. I have of course argued for my own account of objective normative reasons. Rather than relying on the compelling nature of my positive account, however, I shall also argue directly against DBR in section 8.2 below. Consequently, this chapter will be (i) a self-contained challenge to DBR, (ii) part of the supporting structure for my account of the right and the good and (iii) a further reason to think that instrumentalism is an inadequate account of epistemic normativity, and, hence (iv) a further reason to think that naturalism cannot account for value-makers.

The key idea that supports DBR is closely related to one of the main arguments for ethical non-cognitivism. It starts from the observation that norms give us practical and theoretical reasons for action and belief. Moreover, these reasons have a particular sort of power, which we sometimes call normative force. By saying that reasons have normative force we mean that reasons both justify our actions, by standing in their favour, and also that they help to explain why we act as we do, by directly or indirectly motivating us.<sup>3</sup> Thus, the normative force of reasons requires the possibility of coincidence between explanatory and justificatory reasons. How, though, are we to explain this? We might observe that, all things being equal, if one has recognised a normative reason, then to fail to respond appropriately to it is to be practically irrational. Being appropriately responsive to normative reasons is at least part of what it means to be rational. But this is more a statement of the problem, than a solution. For, why should the requirement to be rational have normative force? Although the normative force

---

<sup>3</sup> It is logically possible that there are also reasons which justify, but cannot explain action. For instance, given mind-independence, normative realists must be willing to countenance the possibility of normative reasons that we cannot enter into cognitive contact with or that cannot motivate us. Perhaps, for example, there are justificatory reasons that are beyond human comprehension. Of course, justificatory, but non-explanatory reasons do not constitute an interesting category from the perspective of practical rationality.

of reasons is a familiar part of experience, explaining this phenomenon is one of the more formidable tasks in philosophy.

However we understand the phenomenon of normative force, it places constraints upon what a theory of reasons must be like. Any theory of reasons must be compatible with the fact they make claims on us and have the potential to motivate us.<sup>4</sup> Ideally, a theory of reasons will also *explain* their normative force. For instance, Plato's metaphysics of value attempts to locate the ultimate source of reasons and normativity in the Form of the Good towards which everything strives by its nature. Arguably, the Unmoved Mover of the *Metaphysics* plays a similar role in Aristotle's account of flourishing. Of course, naturalists are disinclined to appeal to spooky entities like Platonic Forms or the Unmoved Mover. However, naturalists also find it difficult to locate reasons anywhere other than in an agent's subjective mental states. For, if we recall the image from chapter one of Hume contemplating a newly disenchanted nature, we will also recall Hume's view that the content of our sensory impressions is always compatible with motivational indifference. We can see that the world is thus and so, and we can simply shrug. Whether or not we are at all motivated by normative reasons depends on our antecedent mental states. In part, this is a matter of what we know about the descriptive facts of the matter. For example, apportioning one's belief to the evidence requires understanding the evidence, and that can require a lengthy process of education. Nevertheless, agents with the same background knowledge with respect to descriptive facts can still respond differently because of differences in character, in virtue, in moral sensitivity and the like.

The fact that external states of affairs are not sufficient to motivate, even defeasibly, should be enough to persuade motivational internalists that reasons must be internal mental states. But it is also likely to persuade naturalists who think that there is *any* reliable connection between normative reasons and

---

<sup>4</sup> To say this is compatible with neutrality regarding the question of whether there is a necessary internal connection between recognising a reason and being motivated to act on it. Although there are various points of contact between the debate over desire-based reasons and the debate over motivational internalism, the two debates are distinct. For instance, one might deny that reasons are mental states whilst maintaining that the recognition of a reason necessarily motivates any (rational) agent to act.



motivation. For, as we have seen in chapter three, the scientific stance aspires to subjectless knowledge and is predicated upon the interchangeability of investigators. However, the fact that only some agents appear to recognise, and to respond to, particular external normative reasons is not fully explained by differences in the agents' background knowledge of scientifically given facts. Along these lines, it is sometimes claimed that the normative implications of any state of affairs are 'essentially contestable'. Consequently, reports of external normative reasons cannot be replicated and verified, and naturalism cannot accommodate them.<sup>5</sup> In the absence of some external state of affairs that would motivate anyone who came into cognitive contact with it, such as the Form of the Good, naturalists are left to focus on the motivational aspects of an agent's psychology. These also appear more scientifically tractable because the explanatory power of normative reasons can then be explained in terms of the causal power of mental states. If the causal powers of mental states are still philosophically puzzling, naturalists will argue that they are at least part of the domain of empirical psychology.

So, we have the following chain of thought. If reasons have normative force, if normative force motivates, and if the only scientifically respectable source of motivation is an agent's mental states, it follows that reasons must depend on mental states. Moreover, it follows from certain widely accepted views about the nature of action and motivation that the relevant mental states must be desires. For, according to the received view, mental states can be divided into beliefs and desires, which jointly explain action. Thus, I might open my fridge because I believe there is fruit juice inside and because I desire to slake my thirst. But, the belief and the desire play different roles in this explanation. Borrowing Anscombe's (1957) helpful and widely used metaphor, beliefs and desires have different 'directions of fit'. Whereas we change the world to fit our desires, we change our beliefs to fit the world. As it is only on the basis of desires that we change the world, it is often maintained that desires are necessary for motivation

---

<sup>5</sup> One could also put the point in terms of scientific 'observation sentences'. Thus, according to Quine: 'An observation sentence is one on which all speakers of the language give the same verdict when given the same concurrent stimulation. To put the point negatively, an observation sentence is one that is not sensitive to differences in past experience within the speech community' (1969a: 86–7). Clearly, on this definition, observation sentences cannot have references to external normative reasons as their content.

and that beliefs are not sufficient.<sup>6</sup> My belief that there is fruit juice inside the fridge will not motivate me to open the fridge in the absence of a desire to slake my thirst.<sup>7</sup> Of course, beliefs and desires are supposed to explain action *jointly*. So, my desire to slake my thirst will not motivate me to open the fridge unless I also believe that there is fruit juice inside. My desire will, however, defeasibly motivate me to find some appropriate means to its satisfaction, whereas the belief in the absence of the relevant desire will not have any particular motivational force. Whereas a belief is merely a vehicle for representing how things stand in the world, a desire is an intrinsically motivating mental state. So, if reasons can motivate us, and if motivation necessarily depends on desires, then reasons must depend on desires and DBR is true.<sup>8</sup>

We have, then, the principal argument in favour of DBR, namely that it is supposed to provide the best explanation of the motivational force of normative reasons. I am not, however, persuaded that the task of explaining normative motivation can be accomplished so easily or that desires provide us with the right sort of basis for normative force. To see why, let us look at the relationship between desire and motivation in a little more detail. Now, it is certainly true that we all have the experience of being motivated by desires. However, the fact that desires can (and regularly do) motivate us does not demonstrate by itself that desires are necessary (or sufficient) for motivation. Indeed, elementary phenomenology suggests that they are not. We regularly experience being motivated by all manner of reasons other than desire. For a start, there are near relatives such as wanting, wishing, yearning, craving, fancying and the like. But

---

<sup>6</sup> This is sometimes called the ‘Humean theory of motivation’ (e.g. Smith 1994; Schroeder 2007). The attribution is based on Hume’s claim that ‘reason alone can never be a motive to any action of the will’ (*T* II, 3, III, 413). For a detailed discussion of the relationship between Hume’s theory of motivation and the ‘Humean theory of motivation’ see Radcliffe (1999).

<sup>7</sup> It is also necessary that belief and the desire are appropriately related or put together. For, I could believe that there is fruit juice in the fridge, desire to slake my thirst, but fail to bring the two mental states together in practical reason. For example, I might not realise that my belief provides me with a way to satisfy my desire, or the fact that I believe that there is fruit juice in the fridge might slip my mind. Schueler (2009) develops this point into an argument against the Humean Theory of Motivation. In order to avoid saying that we only deliberate on the basis of desires – a psychologically implausible claim – Humeans typically argue that desires need only be background conditions, which can be below the level of conscious cognition. However, Schueler argues that in order for us to put beliefs and desires together in practical reason, we must be able to represent desires and that this requirement is in tension with the claim that desires can be subconscious.

<sup>8</sup> Schroeder (2007: 6) calls this the ‘classical argument’ for a Humean theory of reasons.

there are also various motivational influences which run contrary or tangentially to desire. The following examples spring to mind: acting from habit, copying others, acting from grief, from a sense of mischief or devilry, from a sense of duty or obligation, from a spirit of contrariness, from sheer impulse and due to the force of circumstances. When we hear someone saying ‘I desire to  $\phi$ , but intend to  $\psi$ ’ it is not normally a sign of practical irrationality or conceptual confusion, but an indication that there are motivating reasons other than desire.<sup>9</sup>

Although it is difficult to see how these appearances could be denied, Michael Smith (1987: 45–50) has argued against the significance of this sort of elementary phenomenology claim on the grounds that it depends on a mistaken ‘phenomenological conception of desire’. He characterises the mistaken view as follows:

Desires are *like* sensations in that they have phenomenological content essentially, but *differ* from sensations in that they have propositional content as well. (1987: 48)

In order to defeat this view, Smith discusses various examples in which our desires seem to be revealed to us through our observable behaviour rather than being immediately present to consciousness. For instance, someone who avows a desire to be a musician, but takes no steps in that direction, might come to realise that he had no such desire and really just wanted the approval of his music-loving mother. Smith then argues that the fact that we can be deceived about whether or not we have a particular desire is inconsistent with the phenomenological conception of desire, according to which desires have phenomenological content essentially. Smith concludes:

If there is no reason why *any* theorist should accept a phenomenological conception of desire, as we have seen that there is not, then it can hardly be an objection to the Humean’s theory [DBR] that we are unable to introspect the presence of each and every desire that he says we have. (1987: 49)

---

<sup>9</sup> Furthermore, it may well be an indication of being a mature rational agent. See section 8.2 below.

One obvious line of response is to question why Smith forces upon his opponents the view that desires must have phenomenological content *essentially* as opposed to potentially. But, in any case, Smith's conclusion looks too ambitious. Even if it is true that we cannot provide a phenomenological *analysis* of desire, this does not show that there is nothing to be learnt from the phenomenology of desire. The same is true of other mental states such as beliefs, intentions and emotions. There is no general principle to the effect that it must be possible to provide a phenomenological analysis of a mental state in order for phenomenological data to count as evidence concerning the nature of that mental state.

Moreover, the fact that we *can* be mistaken about what motivates us does not show that we are *often* mistaken or that we are deceived in a systematic fashion concerning our motivational influences. Indeed, the very fact that we are sometimes able to step back, reflect and recognise that our desires were not what we took them to be seems to be more a reason for optimism concerning our ability to recognise our motivational influences than a reason to be sceptical about the phenomenology of desire. The objection is not, as Smith suggests, 'that we are unable to introspect the presence of each and every desire' that the proponent of DBR must say we have. For, Smith is right that it is implausible to suppose that we are *infallible* with respect to our own desires. But the concern is that the claim that all of our actions are motivated by desire could only be true if our everyday phenomenology were *radically* mistaken.

Rather than questioning the basic phenomenology of motivation and desire, it is perhaps more likely that an advocate of DBR would accuse me of equivocation. Quite likely it will be pointed out that when it is claimed that desires are necessary for motivation, the term 'desire' is being used in a special sense, such that it is essentially a matter of definition that all voluntary action depends on desire. As a result, it might be objected that my earlier comments on the phenomenology of desire miss the point. A desire, in the sense in which it is often contrasted with a belief, is merely a placeholder for a pro-attitude that disposes an agent to act.<sup>10</sup> Thus, it is trivially true that if an agent hadn't desired

---

<sup>10</sup> See, for instance, Smith (1987: 50–4; 1994: 117) and Schroeder (2007: 146).

to act as he did, he would not have done so; that is assuming that he was not subject to influences beyond his control, in which case the action was involuntary. So, when we hear someone say ‘I desire to  $\phi$ , but intend to  $\psi$ ’, we should interpret this to mean that although they desire to  $\phi$ , their desire to  $\psi$  is stronger. Although it is not entirely clear how we should determine the relative strengths of competing desires, it can be stipulated that our actions reflect the balance of our desires. Indeed, if an agent’s action does not reflect the balance of her desires, there is a sense in which she is not the author of her own actions. Rather she is buffeted by external motivational influences, which are unrelated to the goals and projects that help to make up her identity.

In reply to this, I do not deny that ‘desire’ is sometimes used as a placeholder for a pro-attitude, or that this practice has some advantages. It is worth emphasising, nevertheless, that this is a special, technical usage. Hence, we cannot assume that the sorts of things we would normally say about ‘desires’ in anything like the colloquial sense will apply to ‘desires’ in the technical sense. Indeed, the technical sense is rather opaque insofar as we are still left with the task of distinguishing within the category of desire the various ways in which we are in fact motivated. For, it would appear difficult to say much that is true concerning all the members of the heterogeneous group of desires-as-pro-attitudes. Indeed, it looks like all that holds the category together is the fact that to have a desire that  $p$  is to be disposed to bring it about that  $p$ . If so, however, we cannot rule out the possibility that there could be things which are very unlike desires in the colloquial sense that could count as desires in the technical sense. In particular, for all that has been said, it could turn out that evaluative beliefs or cognitive contact with external reasons, will count as desires in the technical sense.

Smith (1987: 56) argues that there is something incoherent about evaluative beliefs, or ‘quasi-beliefs’ because they would simultaneously have contradictory directions of fit. Insofar as an evaluative belief that  $p$  is a belief it would persist in the face of the perception that  $p$ . Insofar as a desire that  $p$  is motivating, it would not persist in the face of the perception that  $p$ . For the realisation of a desire extinguishes it. Firstly, however, it is far from clear that the metaphor of a

direction of fit is sufficiently robust to support this sort of logical argument.<sup>11</sup> Secondly, we need not allow that an evaluative belief is simply a belief and a desire pasted together in such a way that it inherits the modal properties of its constituent parts. Rather an evaluative belief is the judgement that some state of affairs is a normative reason for action. For instance, it may be the judgment that bringing someone pleasure is good in the circumstances. But, then, there is nothing obviously incoherent about saying that this judgment will motivate us to bring about the good and persist when the good is brought about. Although we will no longer be motivated to bring about the good after we have brought it about, we might well persist in our belief that it is a good.

If it is a merely logical, or definitional constraint that an action must be motivated by a desire, then to deny that evaluative beliefs or external reasons could count as ‘desires’ is to take up a substantive philosophical position. However, there is very strong *prima facie* evidence that evaluative beliefs and external reasons do count as desires in the technical sense. For, when we explain the behaviour of mature rational agents, we are more likely to appeal to objective features of the world or beliefs about what is valuable, than to desires in the colloquial sense.<sup>12</sup> Perhaps, it will be argued that we are radically and systematically deceived in thinking that we are motivated by anything other than our desires. But I do not know how this view could be argued for, if not on the basis that we are *necessarily* motivated by desires. And that is of course precisely the claim in question.

In summary, if the term ‘desire’ is understood in line with common usage, then desires are neither necessary nor sufficient for motivation and we cannot be sure that they will explain why someone acted as he did. It is common for us act on the basis of motivations that are not desires and it is also common for us to act contrary to our desires as a result of competing motivational influences, or as a

---

<sup>11</sup> See Schueler (1991) for development of this sort of criticism.

<sup>12</sup> See Ratcliffe (2007) for a persuasive defence of the view that it is a shared world of norms that enables us to understand and predict each other’s behaviour (rather than belief-desire psychology). He remarks that when we ask why someone acted as they did, ‘what is often expected is a description of a situation that makes clear the relevant norms of activity, rather than an account of people’s psychological predicaments’ (2007: 97).

result of depression or apathy.<sup>13</sup> There is a clear phenomenological licence for these claims and Smith's attempt to undermine it is not compelling. If, on the other hand, desires are understood in the technical desire-as-pro-attitude sense, then it is a substantive question as to whether or not evaluative beliefs or external reasons could fulfil their role. It seems at first blush that they could. At least, in the absence of further independent argument, we are not forced to accept DBR as the only possible explanation of the motivational force of normative reasons.

### *8.2 Desires and the Rationality of Ends*

Although I have offered reasons to doubt that desires are necessary for motivation, I have not denied that desires can and do motivate us. So, for all I have said, it could be that amongst the heterogeneous category of desires-as-pro-attitudes there is a category of mental states which are properly called 'desires' in something like the colloquial sense and transmit normative force from reasons to action or constitute normative reasons in themselves. Although I have argued that DBR is not the only possible explanation of the motivational force of normative reasons, or indeed the most plausible, I have not shown that DBR is false.

In this section, I would like to consider a second line of argument for DBR, which depends upon privileging a particular view of the nature of rationality. Even if desires are not necessary for motivation, it might be argued that it is only when we are motivated to act in the service of our desires that we motivated to act rationally. In other words, it is only in the form of desires that explanatory and justificatory (or motivating and normative) reasons come together. For, it might seem that if we are motivated by anything other than our desires, we are alienated from our ends and, to that extent, less than fully rational, autonomous agents. On this basis, it is sometimes argued that there is nothing more to practical rationality than the rationality one employs in adopting the means to the satisfaction of one's ends. As we shall see below, this appears to have been Hume's view. It also seems to be one of the ideas motivating the kind of instrumentalism we encountered in chapter five.

---

<sup>13</sup> See Cholbi (2011) for more on motivation and depression.

We already have the resources for one response to this line of argument. If desires are understood in the colloquial sense, then it is psychologically and phenomenologically implausible to think that they exhaust our ends. As Schueler writes:

It is difficult to believe that anyone, let alone everyone, reasons on the basis of *nothing but* her own desires all the time. That would be a level of self-absorption almost beyond belief and certainly far beyond the ordinary. (2009: 117)

If, however, desires are just dispositions to achieve whatever ends we have, then DBR is not supported in any interesting sense by the claim that rationality is a matter of successfully pursuing our ends. Nothing substantive has been demonstrated about our ends, and they could be quite unrelated to our desires-in-the-colloquial-sense. For example, they could be objective reasons.

As with the main argument of the previous section, this response is a limited defensive manoeuvre that blunts the argument in favour of DBR. Unless one is persuaded that there is a presumptive argument against DBR, globally or with respect to particular categories of reasons, one would need further arguments in order to adjudicate between DBR and other theories. In this instance, however, I think we can push a little harder against the account of rationality on offer and in so doing apply direct pressure against DBR. Is it really the case that our normative reasons are solely given by what we desire? I shall argue in due course that the answer is ‘no’.

The view that there is nothing more to rationality than instrumental rationality is widespread in economics, game theory and some areas of the social sciences. For instance, the following quotation is from an introductory textbook on decision theory:

Instrumental rationality presupposes that the decision maker has some *aims*, such as becoming rich and famous, or helping as many starving refugees as possible. The aim is external to decision theory, and it is



widely thought that the aim cannot in itself be irrational ... If someone strongly desires to count blades of grass on courthouse lawns, just for the fun of it, that might very well qualify as rational. (Peterson 2009: 5)<sup>14</sup>

Notice the transition from ‘instrumental rationality’ to ‘rationality’ *simpliciter*. Very plausibly, though, the natural response is to object that it *is* irrational to spend one’s time counting blades of grass. Peterson’s reply that ‘everyone should be free to decide for herself what is important in life’ (2009: 5) picks up on the important connection between rational agency and autonomy. Nevertheless, it looks like a weak defence of a life dedicated to counting blades of grass. For, notwithstanding my earlier argument that there is always *some* epistemic value to the possession of true beliefs, it is very hard to see how it could qualify as rational to spend one’s life counting blades of grass when to do so is to incur a very large opportunity cost and be unresponsive to more pressing and important reasons. One cannot make a trivial goal important by an act of will, any more than one can make an immoral goal moral by an act of will. Of course, one can decide to make a trivial goal an important part of one’s life. However, to think that what is an important part of one’s life is important *tout court* would be to have an uncommonly high view of one’s own significance. We might be inclined to think that someone committed to counting blades of grass is a harmless eccentric who should be left in peace, but it is rather pushing it to claim that his activities are important.

The example of counting blades of grass is a relatively benign example of irrationality. Here is Hume in provocative mood:

‘Tis not contrary to reason to prefer the destruction of the whole world to the scratching of my finger. ‘Tis not contrary to reason for me to chuse my total ruin, to prevent the least uneasiness of an *Indian* or person wholly unknown to me. (*T* II, 3, III, 416)

We saw in chapter one how this idea comes from Hume’s naturalism. But, as it is sometimes put, this is the sort of thing that only someone in the grip of a

---

<sup>14</sup> Peterson is of course referring to an example discussed by Rawls (1999: 379–80). Oddly, Peterson thinks that for Rawls a life-plan centred on counting blades of grass is irrational. In fact, Rawls writes that ‘if we allow that his [the grass counter’s] nature is to enjoy his activity, and that there is no feasible way to alter his condition, then surely a rational plan for him will center around this activity’ (1999: 380).

theory could sincerely assert. One feels compelled to respond that it is very evidently contrary to reason to prefer the destruction of the world to scratching one's finger. Indeed, anyone who suggested otherwise would seem to be beyond the pale of rational discourse. Were such a claim sincerely asserted, we would have cause to wonder whether the speaker had lost their grip on reason altogether. They would be a danger to themselves and to others.

Well, Hume may have lost his grip on the concept of reason, but not of course on reason itself. He did not mean to assert that there is no reason to prefer the scratching one's finger to the destruction of the world. Rather, Hume meant that there is no *rational requirement* such that we ought to prefer the one state of affairs to the other. It is our passions, such as benevolence and resentment, which provide us with reasons determine our preferences and they are not, in the main, subject to rational scrutiny.<sup>15</sup> Hence, Hume's famous claim:

Reason is, and ought only to be the slave of the passions, and can never pretend to any other office than to serve and obey them. (*T* II, 3, III, 415)

This is of a course a highly restrictive proposal concerning the scope and nature of reason. Our ends are given by our passions and rational activity is merely a matter of discovering the means to those ends.

So, Humean instrumentalism is a form of anti-rationalism insofar as it denies the possibility of rationally evaluating ends. Although anti-rationalism has gained considerable popularity in recent decades, it is clearly a minority view in the history of philosophy.<sup>16</sup> Hume himself acknowledged that 'this opinion may appear somewhat extraordinary' (*T* II, 3, III, 415). Of course, this is not necessarily a fault and may not concern forward-thinking naturalists.

---

<sup>15</sup> Hume mentions, and then immediately seems to retract, an exception: 'a passion must be accompany'd with some false judgment, in order to its being unreasonable; and even then 'tis not the passion, properly speaking, which is unreasonable, but the judgment' (*T* II, 3, III, 416). Rather optimistically, I think, he continues that 'the moment we perceive the falsehood of any supposition, or the insufficiency of any means our passions yield to our reason without any opposition' (*T* II, 3, III, 416).

<sup>16</sup> As a socio-political aside, it is surely no coincidence that the rise of anti-rationalism in public policy debates has occurred against the backdrop of the failed Soviet experiment. A large part of anti-rationalism's appeal is its apparent support for economic and social liberalism and its opposition to top-down planning and the idea of a common good that overrides the individual good. It also chimes with the more extreme, my-blood-or-thine, libertarian tendency of some contemporary economic thought.

Nevertheless, it does seem that anti-rationalists have some explanatory work to do, perhaps in the form of an error theory. For it has seemed to many philosophers that rational deliberation about ends lies at the very heart of ethics, broadly understood. For example, one might see the central task of Aristotle's *Nicomachean Ethics* as being to rationally evaluate and adjudicate between the lives of hedonism, civic duty and intellectual contemplation; each life governed by a different conception of the end of human flourishing. Moreover, the considerations Aristotle finally presents in favour of a life structured towards the end of intellectual contemplation appear intelligible as rational arguments. On the instrumentalist view of rationality, it is difficult to make sense of this.

The process of moral or normative education is also difficult to make sense of from a Humean perspective.<sup>17</sup> Without doubt, instrumental rationality plays an important role in our lives. But, given their problem-solving abilities, it also seems that something very much like instrumental rationality plays an important role in the lives of nonhuman animals. Certainly, we can use the categories of belief-desire psychology in order to construct plausible explanations of animal behaviour. Thus, we might say that the dog ran towards the tree because it desired to continue chasing the cat and because it believed the cat was in the tree.<sup>18</sup> Moreover, a dog's behaviour is sensitive to its beliefs, or representational states. It may decide not to climb the tree if it realises that the cat it was chasing has now escaped over the rooftops. But, what the dog will not do is stop to consider whether chasing cats is really such a good idea after all. Instead, it will attempt to satisfy whatever desires it happens to have.

The predicament of a human infant seems to be very much like that of many nonhuman animals. At first his behaviour is entirely the product of instinct, but he soon learns that he can modify the world through his behaviour in order to satisfy his desires; that is, he becomes an instrumental reasoner. But, unlike all

---

<sup>17</sup> The structure of the following argument is much influenced by MacIntyre's (1999) book *Dependent, Rational Animals*. See also Hacker-Wright (2009) who emphasises the non-instrumental reasoning of educators and caregivers.

<sup>18</sup> In the absence of a mode of representation for propositional content one might deny that nonhuman animals have beliefs or desires in a strict sense. Be that as it may, they clearly have analogue mental states which enable them to satisfy their goals on the basis of representations of the environments.

or most nonhuman animals, a child also learns that it can be better to modify his desires than to modify the world. Sometimes, this may be a matter of learning that frustrated desires are painful and, so, acquiring a higher-order desire to quell desires that are frequently frustrated. But a child also begins to understand that he is one person among many, and that there are considerations and reasons for action that may be quite unrelated to his desires or immediate viewpoint. Thus, through a process of education and socialisation, children learn to mould and transform their desires in response to the needs and desires of others around them. For instance, Brock and Reader describe the process of moral education thus:

We have to learn that our own needs are not the only ones, still less the only ones that matter, and that our needs exist in a universe full of needs. Learning to recognize and respond to the needs of others is essential if we are to become mature moral agents. (2002: 287)

As I argued in chapter six, a universe full of needs is a universe full of reasons. In time a child's understanding of what counts as a reason expands beyond his own desires and his viewpoint becomes increasingly objective.<sup>19</sup> In particular, he starts to question whether his desires are justified in light of the reasons for action he has learnt to recognise. If he realises that his desires are not justified, he attempts to change them accordingly.

The ability to reason about ends seems distinctively human and it is only when a child learns to do so that it makes sense to think of him as a rational agent. We rely on the instrumental rationality of animals and pre-rational children in order to modify their behaviour with a system of rewards and punishments that correlates with their desires. But, once children learn to question their desires, we can offer them *reasons* to change their behaviour. Some of these reasons will be instrumental to their ends, but many will not be. Moreover, it is only when a child is able to reason about his ends that he becomes responsible for this behaviour. We can reward or punish a pre-rational child or a nonhuman animal, but we cannot properly speaking praise or blame them. If this picture is more or less correct, then the development of instrumental rationality is merely a

---

<sup>19</sup> For a detailed account of child's development in response to the needs of others and her increasingly objective view of moral reasons see Hoffman (2000: Ch.3).

preliminary step along the path to becoming mature rational agents. It is a mistake, therefore, to identify instrumental rationality with rationality *simpliciter*.

Perhaps it will be replied that instrumentalism has the resources to explain what it is to reason about our ends in the only way possible, namely by considering our ends in the light of our existing reasons. As our reasons are given by our desires, this just means bringing our desires into equilibrium. Presumably this would mean making our desires mutually consistent and factoring in their relative strengths. Perhaps, for instance, there is a tension between my desire to spend the afternoon watching amusing videos on YouTube and my desire to spend the afternoon writing an article on metaethics. But, perhaps I also desire to publish a paper on metaethics in order to increase my chances of securing an academic job. Moreover, securing an academic job might fit with a number of my other desires, such as pursuing professional research into philosophy, working in an intellectually stimulating environment and paying the bills. There is, then, an obvious way to bring these desires into equilibrium. If I am rational, and if I am not weak-willed, I will abandon my desire to spend the afternoon watching amusing videos on YouTube. In that way, my remaining desires will be consistent.

But, of course, it is also possible that my over-riding desire is to watch videos of people slipping up on banana skins and that I am willing to structure my life towards the satisfaction of this desire. It might be pointed out to me that I am missing out on far greater goods such as intellectual stimulation, creative enterprise and the dissemination of knowledge. But if I do not desire any of these goods, or if my desire for these goods is much weaker than my desire to watch videos of people falling over, there will be no rational pressure to modify my existing desires. It might even be claimed that it is an advantage of the view in question that it makes my behaviour immune to rational criticism so long as there is no inconsistency among my desires. The thought here would be that as

‘ought’ implies ‘can’, it is not fair to criticise me on the grounds that I am not motivated by desires I do not have.<sup>20</sup>

It strikes me, however, that it is an unattractive consequence of Humean anti-rationalism that we would not be able to provide reasons why someone with an impoverished, but consistent set of desires should broaden their horizons. Far worse, though, is the possibility that my over-riding desire is to act in a way that is immoral. Perhaps, my desire to shoot an innocent man is in tension with my desire to avoid the sight of blood. But perhaps I also desire to own a gun, to run the risk of imprisonment and to revel in the misery of the victim’s relatives. Here, if I am rational, and if I am not weak-willed, it looks like I should overcome my desire to avoid the sight of blood and shoot an innocent man. This is an intolerable conclusion, but I do not see how it can be ruled out if the only constraints upon our reasons are that they are apportioned to the strength of our desires and mutually consistent.

It might be suggested that we should add some further condition to rule out deviant, irrational or immoral desires counting as reasons.<sup>21</sup> Thus, one might appeal to something like what we would desire after undergoing cognitive psychotherapy (Brandt 1979), or what we would desire if we were in possession of all the facts, or in possession of all the relevant facts. However, if we knew all the relevant facts, it is very likely that this would change our view of the world and, so, change our desires. We might then want to appeal to what a fully informed self, or ideal advisor, would desire for us as we currently are. But, of course, our current desires are the product of historically contingent processes some of which (perhaps trauma, or exposure to advertising) may have corrupted us to such an extent that no amount of information or cognitive psychotherapy could cure us of our irrational desires. Perhaps, then, we might want to appeal to what our fully informed self would want for us if we had had a ‘cognitively unimpeachable history’.<sup>22</sup>

---

<sup>20</sup>Shafer-Landau (2003: 181–3) reconstructs this sort of argument on the basis of comments made by Williams (1989). Of course, this line of argument only works on the questionable assumption that desires are necessary for motivation.

<sup>21</sup> See, for example, Smith (1995; 2009) and Schroeder (2007: 83–5).

<sup>22</sup> See Zimmerman (2003: 391) for this proposal. See also Smith (2009: 100–1).

There is an extensive literature on ideal advisors, much of it in relation to preference utilitarianism. Prominent in the literature are proposals for refining the nature of the relationship between our fully informed selves and our real selves, and disputes concerning the appropriate modal semantics (usually cashed out in terms of possible worlds).<sup>23</sup> Without pursuing these subtleties, however, I think it is possible to make three general observations. Firstly, appeals to ideal advisors and the like undermine the *prima facie* attractions of DBR. If my reasons for action are given by the mental states of a fully informed version of myself in another possible world, then the straightforward link between reasons and motivation that DBR appears to provide is lost.<sup>24</sup> Furthermore, the intuitive connection between the satisfaction of desires and autonomous action is also lost. For it is unclear, at best, how my autonomy could consist in acting for reasons grounded in another version of myself with quite different beliefs and desires and located in another possible world. Secondly, there are serious epistemological difficulties. I have no idea what a fully informed version of myself would desire for me, or advise me to do, if not what I would desire for myself after careful deliberation. Again, this is problematic if DBR is supposed to explain motivation. Third, if these refinements preserve the letter of DBR, they do not preserve its spirit. For each refinement is in the direction of an objective theory of reasons. Unless one is persuaded that desires are necessary for motivation, I suggest it is better to bite the bullet and acknowledge that the reason to invoke fully informed selves and the like is to take into account what we have most objective reason to do.

In conclusion, the primary reason to endorse DBR is the thought that desires are necessary for motivation and that this is the only naturalistically acceptable explanation for the phenomenon of normative force. I have argued, however, that there is no compelling reason to think that desires are necessary for motivation. Moreover, I have argued that DBR saddles us with an unattractive anti-rationalism with respect to ends, which is inconsistent with our evaluative

---

<sup>23</sup> For a recent example, see Baber (2010).

<sup>24</sup> There is also the formidable task of providing a naturalistic account of modal facts. For some of the difficulties involved, and an extended argument that naturalism is unable to explain our knowledge of modal properties, see Rea (2002: Ch. 4).

practices. Refinements of DBR to accommodate this tend only to undermine the view's supposed *prima facie* attractions. But, if reasons are not grounded in desires, it might be asked, what grounds their normative force. I made an initial attempt to answer this question at the end of chapter six when I considered the meaning or value of a life as a whole. In the next and final chapter, I shall try a different tack as I also look at questions concerning the epistemology of normative properties.



## Chapter Nine Normativity, Queerness and the Emotions

I would like to conclude the dissertation by arguing for a central role for the emotions in moral epistemology. In the last chapter, I argued that the normative force of reasons need not and should not be explained in terms of desires. In this chapter, I would like to show that the emotions explain how acquiring normative knowledge is reliably, but not necessarily or internally, related to being motivated to act. I also argue that emotional experience provides us with a defeasible warrant for normative knowledge. I frame my discussion in relation to J. L. Mackie's (1977) influential argument from queerness.

### *9.1 Mackie's Argument from Queerness*

Norms have two essential qualities. They are standards against which actions, dispositions of mind and character, states of affairs and so forth can be measured. However, norms also govern our behaviour, make claims on us, bind us and serve as reasons for action with motivational force. It is all very well to grasp the fact that the *telos* of some practice or activity serves as a normative standard against which the excellence of an action can be measured, or that the necessary preconditions for engaging in some practice or activity ground our obligations. However, it is not until one has also grasped the force of a norm *as something to be acted upon* that one can be said to possess normative knowledge in the full sense relevant to practical reason.

Consider J. L. Mackie's well-known argument from queerness, which he summarises as follows:

If there were objective values, then they would be entities or qualities 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 everything else. (1977: 38)

Mackie thinks that the queerness of objective values and the queerness of any putative faculty of moral intuition – the metaphysical and epistemological strands of the argument from queerness respectively – follow from the kinds of facts about normativity just mentioned. In Mackie's words, objective value

would have to be ‘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’ (1977: 40). Such an account of objective value is dramatically exemplified by Plato’s Form of the Good. How then does the queerness of objective value follow from the now familiar thought that a norm is a standard which binds?

Although Mackie does not distinguish them, we can reconstruct his argument in terms of two premises. First, Mackie seems to assume the truth of motivational internalism. Knowledge of the good, he tells us, provides an ‘overriding motive’ such that the knower pursues it.<sup>1</sup> Second, he assumes that motivational internalism is true in virtue of some property of objective goodness. I do not think that we should accept either premise. As I have argued in chapter six and elsewhere, we can make sense of the idea of a person recognising that something is good and not being in the least motivated to pursue it. This is quite clearly true for the functional goodness of artefacts. But it also true for a flourishing life as a whole. Certainly, someone who knows the good he ought to do and doesn’t do it, sins in the sense that he is practically irrational. If, however, we build full rationality into our account of internalism, then internalism is trivially true. For, among other things, rationality is a matter of responsiveness to reasons.

Questions of internalism aside, it does seem clear that there is a reliable and non-accidental connection between recognising a reason for action and acting in accordance with that reason. What I should like to focus on, then, is the second premise. For motivational internalism to be true, Mackie takes it that objective value would need to have ‘to-be-pursuedness somehow built into it’ (1977: 40). It is this property of to-be-pursuedness which appears to makes objective values queer.<sup>2</sup> The property of to-be-pursuedness is likely to seem especially queer if

---

<sup>1</sup> Earlier, Mackie writes that ‘just knowing them or “seeing” them will not merely tell men what to do but will ensure that they do it, overruling any contrary inclinations’ (1977: 23). Here at least Mackie appears to be defending an implausibly strong version of internalism according to which the connection between grasping a norm and being motivated to act in accordance with that norm is non-defeasible.

<sup>2</sup> The tension between moral objectivity and the property of to-be-pursuedness that Mackie highlights is similar to Michael Smith’s (1994) ‘moral problem’ generated by the tension between moral realism and motivational internalism in the light of neo-Humean moral

one takes science to be the ultimate arbiter of ontology. For it might seem unlikely that to-be-pursuedness would feature alongside properties such as mass and charge in a hypothetically completed physics of the future.<sup>3</sup> It is this implausibility that explains the force of the metaphysical strand of the argument from queerness.

In fact, the two strands of his argument, metaphysical and epistemological, are interdependent. The idea that to-be-pursuedness must be built into objective values underpins the epistemological strand of the argument. Mackie elaborates as follows:

When we ask the awkward question, how we can be aware of this authoritative prescriptivity, of the truth of these distinctively ethical premises or of the cogency of this distinctively ethical pattern of reasoning, none of our ordinary accounts of sense perception or the framing and confirming of explanatory hypotheses or inference or logical construction or conceptual analysis, or any construction of these, will provide a satisfactory answer. (1977: 39)

It is difficult to be entirely certain how to interpret this without a clear account of ‘our ordinary accounts of sense perception’. However, it seems likely that Mackie means accounts in terms of the causal interaction between an animal and its environment. The various patterns of reasoning he mentions then operate upon the contents of perceptual experience. The problem, then, is that the property of to-be-pursuedness does not seem to be the sort of property which we can causally interact with (which explains the thought that to-be-pursuedness is unlikely to appear in a final scientific ontology, science concerning itself with causal and constitutive relations). We might say, in a Sellarsian idiom, that to-be-pursuedness is within the space of reasons, not a property causally impinging on that space. Given that to-be-pursuedness is not a property with which we can enter into causal relationships, it is hard to see how we could perceive it.<sup>4</sup> We

---

psychology. This is, Smith claims, ‘the central organizing problem in contemporary meta-ethics’ (1994: 11).

<sup>3</sup> This is the structure of Mackie’s argument, but as I observed in chapter two, who are we to speculate concerning the contents of a hypothetically completed physics of the future? Naturalists should not rule out the existence of particular properties on *a priori* grounds.

<sup>4</sup> The argument has the same formal structure as the standard epistemological argument against Platonism in the philosophy of mathematics (see, e.g., Benacerraf 1973). Mackie’s ‘error theory’ finds its mathematical parallel in Field’s mathematical fictionalism (see, e.g. Field 1989).

might think that we can derive the idea of to-be-pursuedness from the contents of perceptual experience by way of the operations of reason, perhaps akin to the Lockean method of acquiring mathematical knowledge. However, we have already encountered some of the difficulties involved in this approach in chapter one when we studied Hume on 'is' and 'ought'.

In any case, there is a way in which we can undercut this whole line of argument. I observed earlier that Mackie believes that motivational internalism must be grounded in and explained by recognition of the property of to-be-pursuedness somehow built into objective value in the world. However, this does not immediately follow. For it might be that the motivational force of normative structures in the world is grounded in facts about the perception or sensibility of moral agents. It might be, for example, that the faculty of moral perception is such that moral agents are reliably and non-accidentally motivated by normative facts when they perceive them. In this case we would have no need to appeal to a further property of to-be-pursuedness located in the world. So, as a crude first pass, the normative facts are in the world, but the to-be-pursuedness is in the head.

This alternative might not appear live for the moral realist who Mackie is concerned to criticise, because it might seem to relocate an essential feature of normativity away from the objective world and into subjective consciousness. It might be objected, therefore, that a moral realist is committed to the view that norms and, hence, normative force is mind-independent, and that if normative force is mind-independent then the facts which explain an agent's motivation must be found in the external world. However, this line of criticism trades on an ambiguity concerning what it is to 'explain' an agent's motivation. As we saw in chapter four, an explanation might focus either on what it is that *causes* a moral agent to be motivated by his perception of some normative property, or what it is that *justifies* him being thus motivated. Causes and justification can coincide, but they are logically distinct. Moral realism of the sort I have been developing involves the claim that there are objective good-making and right-making properties in the world, and that these properties are such that moral agents are subject to their normative force. Moreover, rational moral agents will be reliably

and non-accidentally motivated when they recognise these normative properties. However, the cause of their motivation need not be some mind-independent property of to-be-pursuedness. The facts that provide the causal explanation of why they are motivated might well be facts about their faculty of moral sensibility. Indeed, this is just the view I shall be arguing for in this chapter.

Moral realism does, however, demand that the *justification* for moral agents being motivated by moral reasons lies in the domain of mind-independent normative facts. Thus, we must explain, in the latter sense, the fact that a moral agent is appropriately motivated in terms of the good-makers and right-makers in the external, mind-independent world. However, there is no need to appeal to some further queer property in the world, a property of to-be-pursuedness beyond the various good-makers and right-makers, in order to explain this. Indeed, such an appeal would be independently implausible. It is not as if the difference in motivational state between Uday, who sees some state of affairs in the world as making moral demands on him, and John, who fails to see that the very same state of affairs makes moral demands on him, is to be explained by the fact that Uday causally interacts with the property of to-be-pursuedness shining over a state of affairs like the star of Bethlehem, whereas John does not. Rather the difference in motivational state is explained by the relevant differences in their respective moral sensibilities. To repeat, however, it does not follow from this that the facts which justify Uday are mind-dependent facts of the sort to trouble a moral realist.

To sum up this section of the discussion, in order to counter the force of Mackie's argument from queerness it would be sufficient to show that moral sensibility has an intrinsically motivational aspect. We could then make sense of the idea that moral agents are justified in being motivated due to the normative facts located firmly in a mind-independent world, whilst what makes it the case that appropriately rational agents will be motivated appropriately are facts about their cognitive capacities. On this view, there is no need to make an appeal to metaphysically queer properties built into the fabric of the world. However, it remains to be shown that we do in fact possess an appropriately structured faculty of moral perception, and that this faculty is not queer.

My view is that we do indeed have just such a faculty of moral perception in the form of emotional experience. The remainder of this chapter will be devoted to defending and elucidating this claim. A reasonably detailed account of the emotions, and more particularly the so-called ‘secondary emotions’, as our faculty of moral sensibility should dispel the impression of queerness.<sup>5</sup> In part this is an empirical claim, and I shall draw on empirical evidence where relevant. However, any such account must be sensitive to the normative phenomena. It needs to be shown not only that we do in fact employ the emotions in moral judgement and perception, but also that the emotions are well suited to pick up on normative facts in the world. Thus, the causal explanation should not undermine the justificatory explanation. I do not pretend to offer anything like a general theory of the emotions, but I shall discuss some of their most important features and argue that these features make emotions cognitive capacities well-suited to detect norms in the world. In particular, I shall argue that emotional experience can be intentional, evaluative and evaluable.

### *9.2 Moral Cognition and the Brain*

At one level, the question of how we go about forming moral judgements is an empirical question; and the empirical evidence strongly supports the thesis that the emotions play a key role. Our ability to reason morally and prudentially, according to social rules, norms and conventions, depends on the proper functioning of the brain’s prefrontal cortices and, in particular, the ventromedial and posterior cingulate regions.<sup>6</sup> In particular, the ventromedial prefrontal cortex (VMPFC – an evolutionarily modern part of the brain located behind the bridge of the nose) appears to be the area of the brain primarily responsible for various ‘secondary’ emotions that play a significant role in moral cognition.<sup>7</sup> This suggests that our ability to engage in moral reasoning and judgement depends upon our ability to exercise certain emotional capacities.

---

<sup>5</sup> I discuss the distinction between ‘primary’ and ‘secondary’ emotions in section 9.4 below.

<sup>6</sup> For a clear overview of the evidence see Greene and Haidt (2002).

<sup>7</sup> These are the so-called ‘secondary’ emotions. I discuss the distinction between primary and secondary emotions below.

Much of the evidence for this claim comes from the work of the neuroscientist Antonio Damasio and his research group (e.g. Saver and Damasio 1991, Damasio 1994). Damasio describes the now famous case of a patient called ‘Elliot’, also referred to as ‘EVR’.<sup>8</sup> Following surgery on a brain tumour, Elliot suffered a lesion to his VMPFC. Elliot’s personal and social life promptly started to disintegrate as he lost his job and family due to increasingly unreasonable behaviour. He was no longer able to guard his own interests or to respond appropriately in social situations.<sup>9</sup> Remarkably, however, he showed no decline in cognitive function as measured by the standard psychological tests. His IQ, short- and long-term memory, production and comprehension of language, powers of facial discrimination, grasp of logical inference and various other cognitive capacities were average or, more often, above average.<sup>10</sup> His emotional capacity, however, was severely affected, as Elliot was himself aware. Damasio describes the change as follows:

Elliot was far more mellow in his emotional display than he had been before his illness. He seemed to approach life on the same neutral note. I never saw a tinge of emotion in my many hours of conversation with him: no sadness, no impatience, no frustration with my incessant and repetitious questioning. I learned that his behaviour was the same in his own daily environment. He tended not to display anger, and, on the rare occasion when he did, the outburst was swift; in no time he would be his usual self, calm and without grudges. (1994: 45)

Despite his emotional indifference and poor social behaviour, Elliot retained his theoretical knowledge of moral, prudential and social principles. Faced with moral and prudential reasoning tests, Elliot’s performance was average or superior. This was in sharp contrast to his real-life performance, which was, as Damasio comments, ‘a catalogue of violations in the domains covered by the problems’ (1994: 46). Elliot’s erratic behaviour was not due to any loss of knowledge, either of normative or non-normative facts. In other words, he knew the better and did the worse. The likely conclusion, then, is that damage to Elliot’s VMPFC damaged his emotional responsiveness to moral and prudential

---

<sup>8</sup> The classic case is that of Phineas Gage (1823–1860) whose prefrontal cortices were destroyed by a tamping iron passing through his skull following an explosion. By all available accounts Gage’s acquired sociopathy was similar to Elliot’s (see A. Damasio 1994: Ch. 1–2; H. Damasio *et al.* 1994).

<sup>9</sup> For a detailed clinical report of Elliot’s behaviour see Eslinger and Damasio (1985).

<sup>10</sup> See Saver and Damasio (1991).

norms, which in turn undermined his motivation to act appropriately. The causal link between his recognition of the to-be-pursuedness of a norm and his motivation to actually pursue it was severed along with his emotional responsiveness. The fact that he remained subject to norms, and his knowledge of that fact, remained intact.

Damasio's original study of Elliot has been extended to twelve other patients with similar prefrontal damage, all of whom show the same combination of poor decision making, flat emotional response and lack of impairment in terms of general cognitive function (Damasio 1994: 53–4). Numerous other studies back up these findings.<sup>11</sup> Thus, Anderson *et al.* note:

It is well established that in adults who have had normal development of social behaviour, damage to certain sectors of prefrontal cortex produces a severe impairment of decision-making and disrupts social behaviour, although the patients so affected preserve intellectual abilities and maintain factual knowledge of social conventions and moral rules. (1999: 1032)<sup>12</sup>

This is all part of a growing body of empirical evidence supporting the thesis that there is an intimate connection between moral judgement and the emotions.<sup>13</sup>

### *9.3 Emotion and Feeling*

So, the empirical evidence cited above strongly suggests a connection between the emotions and our faculty of moral sensibility. Nevertheless, it is far from clear what, if anything, this entails in terms of the moral realism/anti-realism debate. One promising consideration is that emotions have a motivational dimension, or, as psychologists describe it, 'action tendencies'. Emotions involve, as William James maintained, 'an urge to vigorous action' (cited in Solomon 2007: 76), which distinguishes them from cold rational calculation. However, considerable work is still required to show that an emotional basis for

---

<sup>11</sup> See, for example, Price *et al.* (1990), Grattan and Eslinger (1992), Bechara *et al.* (1996), Anderson *et al.* (1999) and Ciaramelli *et al.* 2007.

<sup>12</sup> Anderson's study goes on to provide evidence that damage to the VMPFC in infancy severely impairs the acquisition of moral knowledge, apparently by dramatically reducing responsiveness to punishment. This can lead to extreme forms of anti-social behaviour, which is rare in cases of acquired sociopathy such as Elliot's.

<sup>13</sup> For functional magnetic imagining (fMRI) studies showing that the regions of the brain associated with the emotions are active during moral decision-making see Greene *et al.* (2001), Moll *et al.* (2002) and Greene *et al.* (2004).



moral sensibility can be used to support moral realism against the accusation of epistemological queerness. Indeed, it might seem to suggest the reverse. If moral sensibility depends on the emotions and the emotions are non-rational feelings ill suited to respond to objective normative facts in the world, this would bolster the case for non-cognitivism or an error theory of morality. Non-cognitivism would find support from the idea that moral judgements are expressions of emotional preferences without cognitive content. An error theorist might argue that although moral judgements purport to express truths, they are in fact a product of the emotions which function other than to depict the world as it is and so systematically falsify the nature of reality. For instance, Richard Joyce defends the latter view at (2006: 123–33).<sup>14</sup> If, on the other hand, the emotions have the right properties to function as a faculty of moral sensibility, then we have the promise of a secure account of moral epistemology.

In order to adjudicate between these competing views, we need a better understanding of what an emotion is. The dominant trend in twentieth century psychology was to identify emotions with feelings or unintelligent sensations. In part this was due to a relative lack of interest in the emotions on the part of philosophers, leaving psychologists, and later neuroscientists, to measure what was measurable within a laboratory setting. Thus, the standard scientific view of emotions concentrates on short-term physiological responses, in particular disturbances to the neurological-hormonal-muscular core: facial expression, musculoskeletal responses, effects on the endocrine system and consequent variation in hormone levels, and activation of the autonomic nervous system.<sup>15</sup>

Feeling theories of the emotions have also received some philosophical defence, classically from William James (1884) and more recently from Jesse Prinz (2004; 2005). James states his thesis as follows:

---

<sup>14</sup> See also Joshua Greene, who writes: ‘we believe in moral realism because moral experience has a perceptual phenomenology, and moral experience has a perceptual phenomenology because natural selection has outfitted us with mechanisms for making intuitive, emotion-based judgements, much as it has outfitted us with mechanisms for making intuitive, emotion-based judgements about who among us are the most suitable mates. Therefore, we can understand our inclination towards moral realism not as an insight into the nature of moral truth, but as a by-product of the efficient cognitive processes we use to make moral decisions.’ (2003: 849)

<sup>15</sup> See, for example, Griffiths 1997.

*The bodily changes follow directly the PERCEPTION of the exciting fact, and that our feeling of the same changes as they occur IS the emotion.* (1884: 189–90)

On this view, an emotion just is the perception of bodily disturbance. James' argument is that while a feeling can be accompanied by a judgement or an 'emotional idea' (1884: 196), the two are distinct and it is the feeling that we properly refer to as an 'emotion'. Thus, he writes:

If we fancy some strong emotion, and then try to abstract from our consciousness of it all the feelings of its characteristic bodily symptoms, we find we have nothing left behind, no "mind-stuff" out of which the emotion can be constituted, and that a cold and neutral state of intellectual perception is all that remains (1884: 193)

Again:

What kind of an emotion of fear would be left, if the feelings neither of quickened heart-beats nor of shallow breathing, neither of trembling lips nor of weakened limbs, neither of goose-flesh nor of visceral stirrings, were present, it is quite impossible to think. (1884: 193–4)

On James' view, then, a judgement without an accompanying awareness of physiological responsiveness to the environment is just a judgement, rather than an emotional episode. We are left with a mental state 'purely cognitive in form, pale, colourless, destitute of emotional warmth' (1884: 190). What is distinctive of the emotions is the way they feel.

Although James' thought experiment strongly suggest that emotions are intimately connected with feelings, at least for beings like us, this much is fully consistent with the thesis advanced here that moral sensibility involves the emotions which have an intrinsically motivational dimension. Feelings can be crudely classified into the pleasurable and the painful, and it is little more than a truism to note that we are generally motivated to seek pleasure and to avoid pain. This is so regardless of whether pleasure and pain are our two sovereign masters. The visceral reaction to the perception of an actual state of affairs, or an imagined state of affairs, is crucial to our action tendencies and, thus, feelings are not merely incidental to our moral psychology.

However, James' claim is stronger than this. The argument is intended to show that a judgement without an accompanying awareness of physiological disturbance would not count as an emotion at all. This is certainly not an *a priori* truth. In thought at least, sensations can be detached from the underlying emotional state. It is possible to imagine a different sort of creature, or perhaps a suitably modified or impaired human, who could be angry without the physiological manifestations or sensations. Another putative counterexample is the wrathful Jehovah of the Old Testament as traditionally conceived. In fact, James himself recognises this point, writing:

I do not say that it is a contradiction in the nature of things, or that pure spirits are necessarily condemned to cold intellectual lives; but I say that for *us*, emotion disassociated from all bodily feeling is inconceivable. (1884: 194)

However, there is no need here to resort to controversial thought experiments or the psychology of supernatural beings. In quite ordinary cases, it is clear that certain emotions remains with us long after the violent neurological-hormonal-muscular sensations and the concomitant desire to act die down. Emotions can come in short fiery bursts, but they can also be long-lived, such as the enduring love of a mother for her child or a man for his wife. Of course, there remains some characteristic activity of the nervous system, but, then, the same is presumably true of any mental state. It is implausible, however, to claim that a mother only loves her child when she is conscious of certain accompanying feelings. Minimally, this suggests that there is a distinction to be drawn between Jamesian emotional episodes and our common understanding of the emotions.

There are a number of other reasons for supposing that the Jamesian view of the emotions is incomplete. For example, neither the content of an emotion nor its intentional object is typically a sensation. When I am angry, I am not angry *about* changes in my hormonal balance and the like, but some state of affairs in the world. Moreover, to be the subject of someone else's anger is a matter of concern to us in a way that their physiological states are not. I shall mark out this distinction later in the discussion by referring to 'emotional experience', and outlining in greater detail some of the ways in which emotional experience transcends mere feelings. I contend, then, that there is a close, though

contingent, relationship between emotion and feeling. Although we should reject James' view that an emotion just *is* a feeling caused by a physiological response to some stimulus, we should also acknowledge the pleasurable or painful feelings associated with emotional experience. In the following section, I would like to consider how emotions are more than mere feelings.

#### *9.4 Emotional Experience and Intentionality*

To make matters vivid, and lend the discussion a focal point, let us consider an example of emotional experience described by Peter Goldie:

Imagine you are in a zoo, looking at a gorilla grimly loping from left to right in its cage. You are thinking of the gorilla as dangerous, but you do not feel fear, as it seems to be safely behind bars. Then you see that the door to the cage has been left wide open. Just for a moment, though, you fail to put the two thoughts – the gorilla is dangerous, the cage is open – together. Then, suddenly, you do put them together: now your way of thinking of the gorilla as dangerous is new; now it is dangerous in an emotionally relevant way for you. The earlier thought, naturally expressed as 'That gorilla is dangerous', differs in content from the new thought, although this new thought, thought with emotional feeling, might also be naturally expressed in the same words. Now in feeling fear towards the gorilla you are emotionally engaged with the world, and, typically, you are poised for action in a new way – poised for action out of the emotion. (2000: 61)

Perhaps the first point to notice in this sort of case is that the emotional experience is not a blind sensation. Rather it is intentional in the sense that the emotion is directed towards or about the gorilla. Indeed, the intentional nature of the emotions is commonly remarked on, and the thought can be traced back to Aristotle's *Rhetoric*.<sup>16</sup> To my knowledge, Anthony Kenny (1963) was the first philosopher in the analytic tradition to emphasise the intentional nature of the emotions. Indeed, he argued that emotions are *defined* by their intentional objects and, so, necessarily intentional. For instance, Kenny argued, an emotional attitude counts as fear because it takes something dangerous or fearful as its formal object.<sup>17</sup> Whether or not we endorse Kenny's strong claim that emotions are *necessarily* intentional states, however, it has seemed clear to many philosophers that emotions *can be* intentional states. The main alternative to an

---

<sup>16</sup> See, for example, his discussion of anger (1379a–1379b).

<sup>17</sup> Solomon also argues that emotions are necessarily intentional: 'no feeling and no physiological response even counts as emotional unless it has the property of intentionality' (2007: 205). See also Sartre (1962: 57).

intentional view of emotional experience is a bare Jamesian view which we have already rejected.

Following from the thought that emotions have intentional content, we might notice that in emotional experience we take the world to be a certain way. Emotional experience purports to attune us to the world and focus our attention on its salient features. To be fearful of a gorilla is, in part, to take it that the gorilla is dangerous. In this sense, an emotion is more akin to a belief than a desire. In Anscombe's terms, emotional experience has a mind-to-world direction of fit. Consequently, there is something that it is for emotions to be appropriate or fitting and this depends on how the world is. Emotional experience depicts the world in a certain way, and gets it right when the world is as depicted in emotional experience. To be fearful of an escaped gorilla may be appropriate. To be fearful of a gorilla inside a secure cage is a mistake.

Emotional experience can get things right in another way. Whereas a belief gets things right when its propositional content is true, emotional experience can also be evaluated in terms of its phenomenal intensity. If I am right that emotional experience sometimes serves as the vehicle for normative judgments, this reflects the fact that value comes in degrees. So, sometimes our emotional experience has the appropriate direction, but not the appropriate intensity. It would be difficult to make sense of someone both expressing a high degree of moral indignation and simultaneously accepting that no great wrong has occurred. The recognition that the emotional response is out of proportion to the offence may in itself calm the emotional response. But it should at least prompt recognition that the emotional response is irrational or inappropriate. Alternatively it might lead to *post hoc* rationalisations or confabulations in order to avoid cognitive dissonance; for example a search for reasons that justify the initial violence of the emotion. Whilst this is irrational, it is only intelligible on the basis that emotional intensity is subject to rational evaluation.

Not only are occurrent emotions subject to rational evaluation, so are emotional dispositions. I discussed doxastic voluntarism in chapter six. As I am arguing that emotions are belief-like in some respects, we might expect that the broad

outlines of that discussion will also apply here. And, indeed, that is the case. Generally speaking, our emotions are not states that happen to us or with which we are afflicted. Although we are occasionally overcome by emotion, this is the exception rather than the rule. Moreover, whether we are overcome in the moment normally depends on our background beliefs and dispositions. As with beliefs, we can cultivate our emotional dispositions directly or indirectly. We can realise that we are too quick to anger, and learn to count to ten or redirect our attention. We can make efforts to sympathetically imagine the situations of others. By listening to people, we can learn to respond appropriately to morally salient facts that we might not recognise from our own perspective. We are also inveterately story-telling beings and we learn from each other's tales of everyday life and of make-believe. Sometimes, we refine our emotional dispositions by paying attention to subtle moral distinctions in plays, novels, films and other narrative arts. All of this tends to produce virtuous emotional dispositions.<sup>18</sup>

On the other hand, we can cultivate emotional experience in a way that does not reflect normative reality. For instance, Arlie Hochschild (1983) describes self-induced emotions in airline stewardesses who are expected to cultivate a general disposition of cheerfulness even when it is not the emotional response most appropriate to the situation.<sup>19</sup> They might have followed William James' advice:

Smooth the brow, brighten the eye, contract the dorsal rather than the ventral aspect of the frame, and speak in a major key, pass the genial compliment and your heart must be frigid indeed if it does not gradually thaw. (1884: 198)

More crudely, we can ply ourselves with excessive alcohol, or listen to stirring music, in order to intensify feelings of joyfulness, bravery or melancholy. Partly because emotional experience has a felt dimension, the cultivation of intense emotions can be pleasurable. On the other hand, we can become addicted to the felt aspects of emotional experience and that can be a severe harm. From the point of view of normative epistemology, however, manipulating one's

---

<sup>18</sup> See Brady (2010: 117f.) for various other ways in which we pay 'virtuous attention to our emotional systems'.

<sup>19</sup> See also Solomon (2007: 197–8).

emotional states in such a way that they become less reliable guides to normative reality is always a vice.

So, I suggest that the content and intensity of emotional experience can be rationally evaluated in terms of their fittingness to the world and that we are subject to evaluation on the basis of the emotional dispositions we cultivate. I would now like to argue that in emotional experience the world is disclosed to us in a new way. In particular, emotional experience is a mode of representation that reveals normative properties as salient. Thus, in the course of his phenomenological enquiry, *Sketch for a Theory of the Emotions*, Sartre calls an emotion ‘a transformation of the world’ (1962: 63). In an attractive, but rather opaque phrase, he also talks about an emotion as ‘a sudden fall of consciousness into magic’ (1962: 90). Expanding on the metaphor of magic he writes:

We have seen how, during an emotion, the consciousness abases itself and abruptly transmutes the determinist world in which we live, into a magical world. But, conversely, sometimes it is this world that reveals itself to consciousness as magical just where we expect it to be deterministic. It must not, indeed, be supposed that magic is an ephemeral quality that we impose upon the world according to our humour. There is an existential structure of the world which is magical. (1962: 84)

Although the idea is not precise, Sartre’s idea of magic seems to be as sort of post-Weberian re-enchantment of the world that makes some courses of action, but not others, appear possible. He largely focuses on cases in which we project our emotions on to the world in order to excuse ourselves from acting and to reduce cognitive dissonance.<sup>20</sup> As these are mere projections of magic, Sartre seems for the most part to endorse something like an error theory of the intentional content of emotions. So, it is far from clear that he is entitled to say, as he does in the above quotation, that the world reveals itself as magical or that it has a magical existential structure.<sup>21</sup> Regardless of whether Sartre was consistent, however, the above quotation suggests an attractive account of the emotions as modes of representation, like perceptions, with intentional content, that are distinct from scientific (deterministic) representations and reveal the

---

<sup>20</sup> See, for example, Sartre’s discussion of the girl who breaks down in emotion because she cannot face the prospect of caring for her sick father (1962: 31f.).

<sup>21</sup> Weberman (1998: 396) also presses this line of objection.

world as having particular ‘magical’ properties that are significant for action (have normative force).

I think this view of the emotions is approximately correct, although I do not think that we need to follow Sartre in thinking that value in the world is magical. In fact, I have already provided a non-magical account of normative properties in chapter six. Nevertheless, I think it is right that in emotional experience, we have a quasi-perceptual understanding of the world as charged with normative force and significance. I shall return to this idea. First, though, I want to make clear that I am not attempting to provide an *analysis* of emotional experience. The aspects of emotional experience with which I am particularly concerned do not characterise all emotions. For, emotions are a heterogeneous kind and there are a number of clear counterexamples to my central claims that emotional experience is intentional, subject to rational evaluation, evaluative and analogous to perception.<sup>22</sup> These exceptions fall into two categories and help us to refine our view.

The first exceptions are primary emotions, including certain primitive forms of anger, fear and disgust. Primary emotions are pre-cognitive, more or less hard-wired responses to environmental stimuli. In this sense they are like the startle response. They are pan-cultural and also present in many other animals.<sup>23</sup> So, we can sometimes speak of a person being angry in much the same way that a cat is angry. Whereas a cat raises its hackles and flashes its tail, a person turns red and grimaces to bare his teeth. In both cases, they are signalling that they are prepared to fight and the bodily manifestations of the emotion prime them to do so. From a physical perspective, primary emotions are correlated with activity in the amygdala, part of the evolutionarily old limbic system.<sup>24</sup> By contrast, secondary emotions arise from an interaction between the limbic system and the

---

<sup>22</sup> Paul Griffiths (1997: Ch. 9) argues that there is little explanatory power and no theoretical unity to the category of emotions, and that consequently, the term ‘emotion’ does not pick out a natural kind and should be eliminated from theoretical discourse.

<sup>23</sup> See Ekman (1992) for evidence of the pan-cultural nature of primary emotions.

<sup>24</sup> For detail on the role of the amygdala on moral and social behaviour, see Adolphs (1999). See also Greene and Haidt (2002: 522).



prefrontal cortices.<sup>25</sup> Whereas the neural substrate of primary emotions is entirely in the subcortex, the secondary emotions involve the subcortex and neocortex working together.<sup>26</sup> Although all emotions recruit and depend on the feelings which constitute primary emotions, the sort of emotional experience I am interested in here involves the combination of Jamesian feeling and representational content that is typical of the secondary emotions.<sup>27</sup> Goldie captures something of this combination with his description of emotions as ‘feeling towards’ or, equivalently, ‘thinking of with feeling’ (2000: 58). Notice also, that the secondary emotions can be used to evaluate the primary emotions and other secondary emotions. So, I can be proud of my courage or ashamed of my lack of empathy. I can also be frightened of my fear, which shows that some emotions come in primary and secondary forms. The fear I am frightened of might be a primary emotion, whereas the fright I feel of my fear is a secondary emotion because it has intentional content. This tends to confirm the idea that secondary emotions are evaluative.

Very generalised emotions, such as moods, form another category of emotions that do not seem to be intentional. Thus, in arguing against cognitive account of the emotions, Paul Griffiths (1997: 28) objects that emotions such as depression, elation and anxiety do not have a clear object. Sartre anticipates this sort of objection and responds as follows:

Even if it is a case of one of those indefinite anxieties that one feels in the dark in a sinister and deserted passage, etc., it is still *of* certain aspects of the night, or of the world, that one is afraid. (1962: 57)

Perhaps Sartre is right that emotional states such as *angst* or *joie de vivre* are prompted by features of the world. We are certainly more likely to enjoy a sense of *joie de vivre* on a warm, bright summer’s than on a drab winter’s evening. However, it is not always true that the cause of an emotion figures in its intentional content. For example, I might feel anxious because I have drunk too

---

<sup>25</sup> For details on the distinction between primary and secondary emotions and an explanation of the underlying physical mechanism see Damasio (1994: 131–42).

<sup>26</sup> In particular, the central cingulate is known to integrate basic emotions and cognition. See Panksepp (2003: 238).

<sup>27</sup> As Damasio writes: ‘It is the connection between an intricate cognitive content and a variation on a preorganized body-state profile that allows us to experience shades of remorse, embarrassment, *Schadenfreude*, vindication, and so on’ (1994: 150).

much coffee. It does not follow that I am anxious about the coffee. Similarly, it does not follow from the fact that my anxiety is caused by ‘certain aspects of the night’ that I am anxious of those aspects. Moreover, in this particular case it is not plausible that our ‘indefinite anxieties’ *must* have ‘aspects of the night’ as their intentional objects. If indeed they did, they would not be so ‘indefinite’. For this reason, I propose to exclude existential moods and highly generalised emotions, as well as basic emotions, from my discussion of emotional experience.<sup>28</sup>

### 9.5 *Emotion and Perception*

Returning to our main theme, I suggested above that emotional experience is quasi-perceptual. In particular, it is a forming of *perceiving as*. In emotional experience, we perceive a gorilla *as* dangerous or we see an act *as* contemptible. In one way, this is just as we perceive an apparently elliptical piece of engraved copper as a round penny. Our experience depends on our virtues and vices, on our stable dispositions of character and thought. However, in emotional experience, we do not come to see the world in a new way in virtue of perceiving something beyond its sensible or perceptible properties. In Goldie’s example, the difference between before and after putting the two thoughts together – gorilla dangerous, cage open – is not that one has literally seen a new property of the gorilla, the property of to-be-fearedness. There is no skywriting in the new representation of the gorilla with the word ‘dangerous’ scrawled across the top.<sup>29</sup> The gorilla’s sensible properties are the same. It is to be feared because of properties such as its weight, its strength and its propensity for aggression in certain circumstance. These are all scientifically available, natural properties.

In the quotation at the start of this section, Goldie related this new way of seeing the world to the issue of motivation. As he wrote:

---

<sup>28</sup> In fact, I am inclined towards the view that existential moods can be judgements concerning how one’s life is going as a whole and, so, related to the concerns about life’s meaning or value mooted in chapter six. I hope to develop this idea elsewhere.

<sup>29</sup> This lack of representational content is the explanation I offered of Hume’s discussion of ‘is’ and ‘ought’. See Millgram’s (2005: 235) comment on ‘skywriting’ that I quoted in chapter one.

In feeling fear towards the gorilla you are emotionally engaged with the world, and, typically, you are poised for action in a new way – poised for action out of the emotion. (2000: 61)

Like zoologists, we can coolly perceive the various sensible properties in virtue of which a gorilla is to be feared. However, this cool perception does not determine that we will fear the gorilla or that we will be motivated to act in virtue of the danger it poses to us. Rather it is when we perceive the gorilla as dangerous *in emotional experience* that we are poised to act; hence, Sartre's idea of transmuting 'the determinist world in which we live, into a magical world' (1962: 84). It may seem no surprise that emotional experience tends to produce motivation. For, as I commented in section 9.3 above, emotional experience has a felt dimension that can be pleasurable or painful. But whilst this is evidently a spur to action, I think it is only a partial explanation. More important is Goldie's thought that in feeling fear we are emotionally engaged with the world. Consequently, we recognise normative structures in the world as salient and as constituting reasons for *us*. Indeed, I suggest that the recognition of reasons in emotional experience is a significant motivating element in our moral psychology. In this sense, the passions are the servants of reason. We are rational to the extent that we act on the basis of reasons, and it is in emotional experience that reasons become salient for *us* in light of our practical and theoretical commitments.

Perhaps my account of the role of the emotions in normative cognition will seem like grist for the mill of the anti-realist. Indeed, as I mentioned above, Joyce (2006) defends an error theory of morality partly on the basis of the prominent role the emotions play in moral judgement. I anticipate the reasonable concern that perceiving as or seeing as, unlike perceiving or seeing, is not necessarily factive. If I see a rabbit in a field, then, *ceteris paribus*, I know that there is a rabbit in the field. But, if I see something in a field *as* a rabbit, it does not follow, *ceteris paribus*, that I know there is a rabbit in the field. I might well be seeing a hare or a shadow. Of course, we can also talk about "seeing things *as* they are". So, perhaps it would be a mistake to read too much into these locutions. Nevertheless, the concern remains that even if the world is represented as having particular normative properties in emotional experience,

the emotions are not *reliable* cognitive capacities. Interestingly in light of my criticisms of epistemic instrumentalism in chapter five, the main reason for this concern might be that emotional experience is targeted at action not truth.

In fact, I do think that there are good reasons to be cautious about the reliability of our emotions as cognitive capacities. As we have seen, emotions function as action tendencies and as coping mechanisms. Moreover, if emotions are intentional, we can be wrong about the object of our emotions. As Freud observed, emotions can be displaced. Similarly, emotions can spread out beyond their appropriate object. For example, disappointment concerning some particular failure or injustice can transmogrify into depression and poison our view of life and people'. Alternatively, a joyful event can cause us to view the world through 'rose-tinted spectacles'. Moreover, as I have mentioned, emotions can be directly manipulated for non-rational ends.

In response to these sorts of worries, I would like to make three brief, initial points. First, these phenomena support the claim that emotions do have an object, and that the relationship between emotion and object can be more or less appropriate. So, even if emotional experience sometimes misrepresents the world, this is against the background expectation that emotional experience gets things right. Second, if we accept that emotional experience represents the world as to-be-acted-upon, there are (contingent) evolutionary reasons to think that they will be reliable more often than not. These are the same considerations instrumentalists adduce in favour of the idea that epistemic norms can be reduced to pragmatic norms. Third, we are often able to recognise when the normative structure of the world is misrepresented in emotional experience. Sometimes, as I have discussed, this is a matter of bringing our emotions to bear on one another. This is no more worrying in principle than correcting perception with perception. However, it also suggests an independent standard of normative appraisal – grist for the realist's mill.

The most important point, however, is that I am not claiming that emotional experience constitutes the grounds for normative judgments.<sup>30</sup> Rather, I am arguing that emotional experience provides a defeasible warrant for normative judgments and also motivates us in virtue of the way that reasons became salient for us. I am happy, therefore, to allow that there is an independent standard against which emotional content can be evaluated. Indeed, this is a necessary aspect of my overall account. Emotional experience should be responsive to the kind of normative reasons I identified in terms of the right-makers and good-makers for the contingent practices and non-contingent activities of beings like us. Moreover, I do not claim that there is a necessary connection between moral judgment and emotional experience, or between moral motivation and emotional experience. What I should like to argue, and what is necessary in order to address Mackie's argument from queerness, is that moral agents have a faculty of moral perception in virtue of which they are reliably and non-accidentally motivated by normative facts when they perceive them.

### 9.6 Conclusion

Quine talks about 'the deep old duality of thought and feeling, of the head and the heart, the cortex and the thalamus, the words and the music' (1981: 55). But I have argued in this chapter that the duality is not as deep as all that. In normative cognition, the two pull together as the heart works in service of the head by drawing our attention to normatively salient facts. When, in emotional experience, we perceive something as a normative reason for action that is defeasible evidence for the truth of the corresponding normative proposition. Moreover, it is an explanation of the fact that we are often motivated by normative considerations.

Appreciating the role of the emotions in normative cognition moves us further along in understanding why value-makers present a problem for naturalism. Whilst emotional experience helps to explain both normative knowledge and motivation, references to the emotions are significantly absent from the context

---

<sup>30</sup> Except in cases where *the fact that* someone is in a particular emotional state constitutes a reason for action.

of scientific justification. This is partly because emotions *can* distort perception and partly because emotional experience is not quantifiable. More important, however, is the fact that the content of emotional experience depends on a subject's perspective, dispositions and practical commitments. As we saw in Goldie's example of the escaped gorilla, emotional experience represents aspects of the world as reasons for *us* to act. Thus, emotional experience is subjective in the sense that contrasts with science's regulative ideal of subjectless knowledge. As we have also seen, however, this does not mean that the intentional content of emotional experience is subjective in the sense that it depends for its truth on the mental states of the subject. A gorilla is dangerous because of its common or garden physical properties and its escape is a reason for us to act because of our non-contingent needs.

The account of normativity I have defended in the second half of this dissertation might be considered naturalistic in the weak sense that it involves no eldritch metaphysical claims or appeals to supernatural agency. Moreover, it is not, I hope, incompatible with our best science. It is, however, incompatible with scientism. The methods I have used to establish and justify this account have been drawn from the full range of the traditional philosophical armoury. For example, *a priori* argument, intuition, emotional experience, phenomenological investigation, linguistic analysis and conceptual analysis have been brought to bear alongside empirical observation, deductive and inductive logic. In fact, I have reached my conclusions in the standard philosophical way, by marshalling and reflecting on whatever sources of evidence seem relevant to understanding the general structure of things.

I argued in chapter two that the best way to avoid a dialectical impasse between epistemological naturalists and the traditional mainstream in philosophy is to judge the respective research programmes by their fruits. In chapters four and five, I identified some of the weaknesses and limitations of naturalistic approaches to normativity. By contrast, I have attempted to demonstrate that by enlarging our repertoire of methods, we are able to develop a plausible and attractive account of the nature of normativity. Thus, I have defended an account of what normative properties are, how we know about them and how they

motivate us to act. In other words, I have defended, in broad outline, a non-naturalistic solution to the problem of normativity.

## **Bibliography**

- Aach, J. 1990. Psychologism reconsidered: a re-evaluation of the arguments of Frege and Husserl. *Synthese* 85: 315–38.
- Adolphs, R. 1999. Social cognition and the human brain. *Trends in Cognitive Sciences* 3: 469–79.
- Alcázar, J. R. 1993. Epistemic aims and values in W. V. Quine's naturalized epistemology. *Philosophical Issues* 3: 309–318.
- Almeder, R. 1998. *Harmless Naturalism*. Chicago: Open Court.
- Alston, W. 1986. Internalism and externalism in epistemology. *Philosophical Topics* 14: 179–221.
- Alston, W. 1988. The deontological conception of epistemic justification. *Philosophical Perspectives* 2: 257–99.
- Alston, W. 2005. *Beyond Justification*. Ithica, N.Y.: Cornell University Press.
- Anderson, S., Bechara A, Damasio H., Tranel, D., Damasio, A. 1999. Impairment of social and moral behaviour related to early damage in human prefrontal cortex. *Nature Neuroscience* 2: 1032–1037.
- Anscombe, G. E. M. 1957. *Intention*. Oxford: Basil Blackwell.
- Aristotle. 1999. *Nicomachean Ethics*. Trans. T. Irwin. Cambridge, Indianapolis: Hackett Publishing.
- Armstrong, D. M. 1980. Naturalism, materialism and first philosophy. In his *The Nature of Mind*. Brighton: The Harvester Press, 149–65.
- Armstrong, D. M. 1997. *A World of States of Affairs*. Cambridge: Cambridge University Press.
- Austin, J. L. 1962. *Sense and Sensibilia*. Oxford: Oxford University Press.
- Baber, H. E. 2010. Worlds, capabilities and well-being. *Ethical Theory and Moral Practice* 13: 377–92.
- Baillie, J. 2000. *Hume on Morality*. London: Routledge.
- Baldwin, T. 1990. *G. E. Moore*. London: Routledge.
- Bechara, A, Tranel D, Damasio H, Damasio A R. 1996. Failure to respond anatomically to anticipated future outcomes following damage to prefrontal cortex. *Cerebral Cortex* 6: 215–225
- Beilby, J. (ed.). 2002. *Naturalism Defeated*. Ithaca: Cornell University Press.
- Benacerraf, P. 1973. Mathematical truth. *Journal of Philosophy* 70: 661–79.
- Bennett, J. 1964. *Rationality: An Essay Towards Analysis*. London: Routledge and Kegan Paul.
- Bennett, J. 1990. Why is belief involuntary? *Analysis* 50: 87–107.
- Berger, A. 2003. The Quinean quandary and the indispensability of nonnaturalized epistemology. *The Philosophical Forum* 34: 367–81.
- Bigelow, J. 1988. *The Reality of Numbers*. Oxford: Oxford University Press.
- Blachowicz, J. 2009. How science textbooks treat scientific method: a philosopher's perspective. *British Journal for the Philosophy of Science* 60: 303–44.
- Blackburn, S. 1984. The individual strikes back. *Synthese* 58: 281–301. Reprinted in Miller and Wright, eds. (2002), 28–44.
- Blackburn, S. 1998. *Ruling Passions*. Oxford: Clarendon Press.
- Bloomfield, P. 2006. Opening questions, following rules. In Horgan and Timmons (eds.), 2006, 169–89.
- Boghossian, P. 1989. The rule-following considerations. *Mind* 98: 507–49. Reprinted in Miller and Wright, eds. (2002), 141–87.



- Booth, A. R. 2007. Doxastic voluntarism and self-deception. *Disputatio* 22: 115–30.
- Brady, M. 2010. Virtue, emotion, and attention. *Metaphilosophy* 41: 115–31.
- Brandt, R. 1979. *A Theory of the Good and the Right*. Oxford: Oxford University Press.
- Brink, D. 1989. *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge University Press.
- Brock, G. and S. Reader. 2002. Needs-centred ethical theory. *The Journal of Value Inquiry* 36: 425–34.
- Brown, H. I. 1988. Normative epistemology and naturalized epistemology. *Inquiry* 31: 53–78.
- Bykvist, K and A. Hattiangadi. 2007. Does thought imply ought? *Analysis* 67: 277–285.
- Bruening, W. 1971. Moore on ‘is-ought’. *Ethics* 81: 143–9.
- Camus, A. 1955. *The Myth of Sisyphus*. London: Penguin.
- Card, R. F. 2007. Inconsistency and the theoretical commitments of Hooker’s rule-consequentialism. *Utilitas* 19: 243–58.
- Carnap, R. 1967. *The Logical Structure of the World*. Berkeley: University of California Press.
- Caro, M. and D. MacArthur (eds). 2004. *Naturalism in Question*, Cambridge, Mass.: Harvard University Press.
- Carroll, L. 1895. What the tortoise said to Achilles. *Mind* 4: 278–80.
- Carruthers, P., S. Stich and M. Siegal (eds.). 2002. *The Cognitive Basis of Science*. Cambridge: Cambridge University Press.
- Cassam, Q. 2011. Knowing what you believe. *Proceedings of the Aristotelian Society* 111: 1–23.
- Casebeer, W. D. 2003. *Natural Ethical Facts*. Cambridge, Mass.: MIT Press.
- Chalmers, A. F. 1999. *What is This Thing Called Science?* 3<sup>rd</sup> edition., Buckingham: Open University Press.
- Chase, W. and H. Simon. 1973. Perception in chess. *Cognitive Psychology* 4: 55–81.
- Chisholm, R. M. 1991. Firth and the ethics of belief. *Philosophy and Phenomenological Research* 51: 119–28.
- Cholbi, M. 2011. Depression, listlessness and moral motivation. *Ratio* 25: 28–45.
- Churchland, P. 1987. Epistemology in the age of neuroscience. *The Journal of Philosophy* 84: 544–552.
- Ciaramelli, E., M. Muccioli, E. Ladavas and G. di Pellegrino. 2007. Selective deficit in personal moral judgement following damage to ventromedial prefrontal cortex. *Social Cognitive and Affective Neuroscience* 2: 84–92.
- Clifford, W. K. 1877. The ethics of belief. In his *The Ethics of Belief and Other Essays*, Prometheus Books, 1999.
- Cohen, G. A. 2000. *If You’re an Egalitarian, How Come You’re So Rich?* Cambridge, Mass.: Harvard University Press.
- Colyvan, M. 2001. *The Indispensability of Mathematics*. Oxford: Oxford University Press.
- Cottingham, J. 2003. *On The Meaning of Life*. London: Routledge.
- Crane, T. 1993. Reply to Pettit. *Analysis* 53: 224–7.
- Crane, T. and D. H. Mellor. 1990. There is no question of physicalism. *Mind* 99: 185–206.

- Damasio, A. 1994. *Descartes' Error*. London: Vintage.
- Damasio, A. R. 1997. Towards a neuropathology of emotion and mood. *Nature* 386: 769-770.
- Damasio, H., T. Grabowski, R. Frank, A. M. Galaburda and A. Damasio. 1994. The return of Phineas Gage: clues about the brain from the skull of a famous patient. *Science* 264: 1102-5.
- Dancy, J. 2000. *Practical Reality*. Oxford: Clarendon Press.
- Darwall, S., A. Gibbard and P. Railton. 1992. Toward *fin de siècle* ethics: some trends. Reprinted in their (eds.), 1997, *Moral Discourse and Practice*, New York: Oxford University Press, 3-47.
- Davies, P. S. 2001. *Norms of Nature*. Cambridge, Mass.: MIT Press.
- Dawkins, R. 1999. The value of science and the science of values. In W. Williams (ed.) *The Values of Science*, Boulder, Colorado: Westview Press, 11-41.
- Day, R. 1979. *How to Write and Publish a Scientific Paper*. Philadelphia: Institute for Scientific Information Press.
- Dennett, D. 1999. Faith in the truth. In W. Williams (ed.) *The Values of Science*, Boulder, Colorado: Westview Press, 95-109.
- Dennett, D. 2006. *Breaking the Spell*. London: Allen Lane.
- Devitt, M. 1998. Naturalism and the a priori. *Philosophical Studies* 92: 45-65.
- Dewey, J. 1929. *Experience and Nature*. New York: Dover.
- Dummett, M. 1959. Truth. *Proceedings of the Aristotelian Society* 59: 141-62.
- Dunbar, K. N. 2002. Understanding the role of cognition in science: the *science as category* framework. In Carruthers et. al. (eds.) 2002, *The Cognitive Basis of Science*, 154-70.
- Ekman, P. 1992. Facial expressions of emotion: New findings, new questions. *Psychological Science* 3: 34-8.
- Engel, P. 2005. Logical reasons. *Philosophical Explorations* 8: 21-38.
- Eslinger, P. and Damasio, A. 1985. Severe disturbance of higher cognition after bilateral frontal lobe ablation: patient EVR. *Neurology* 35: 1731-41.
- Evans, G. 1982. *The Varieties of Reference*. Oxford: Clarendon Press.
- Evans, J. L. 1962. Grade not. *Philosophy* 37: 25-36.
- Feldman, R. 2000. The ethics of belief. *Philosophy and Phenomenological Research* 60: 667-95.
- Feyerabend, P. 1975. *Against Method*. London: Verso.
- Field, H. 1989. *Realism, Mathematics and Modality*. Oxford: Basil Blackwell.
- Finlay, S. 2007. Four faces of moral realism. *Philosophy Compass* 2: 1-30.
- Flanagan, O. 2007. *The Really Hard Problem: Meaning in a Material World*. Cambridge, Mass.: MIT Press.
- Fogelin, R. 1997. Quine's limited naturalism. *The Journal of Philosophy* 94: 543-63.
- Foot, P. 1972. Reasons for action and desires. In her *Virtues and Vices*, Oxford: Basil Blackwell, 148-56.
- Foot, P. 2001. *Natural Goodness*. Oxford: OUP.
- Forbes, G. 1984. Scepticism and semantic knowledge. *Proceedings of the Aristotelian Society* 84: 223-37. Reprinted in Miller and Wrights, eds. (2002), 16-27.
- Frankena, W. 1939. The naturalistic fallacy. *Mind* 48: 464-77.
- Frankfurt, H. 1965. Descartes' validation of reason. *American Philosophical*

- Quarterly* 2: 149–56.
- Frankfurt, H. 1984. Necessity and desire. *Philosophy and Phenomenological Research* 45: 1–13.
- Frege, G. 1893 / 1953. *The Foundations of Arithmetic*, trans. J. L. Austin. Oxford: Blackwell.
- Frege, G. 1897. Logic. In his *Posthumous Writings*, trans. P. Long et al. Chicago: University of Chicago Press, 127–51.
- Fricker, M. 2007. *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford: Oxford University Press.
- Geach, P. T. 1956. Good and evil. *Analysis* 17: 33–42.
- Geach, P. T. 1977. *The Virtues*. Cambridge: Cambridge University Press.
- Giere, R. 1987. The cognitive study of science. In N. J. Nersessian (ed.) *The Process of Science*, Dordrecht: Martinus Nijhot, 139–59.
- Giere, R. 1989. Scientific rationality as instrumental rationality. *Studies in History and Philosophy of Science* 20: 377–84.
- Glassen, P. 1957. A fallacy in Aristotle’s argument about the good. *Philosophical Quarterly* 7: 319–22.
- Godden, D. M. 2005. Psychologism in the logic of John Stuart Mill: Mill on the subject matter and foundations of ratiocinative logic. *History and Philosophy of Logic* 26: 115–43.
- Goldie, P. 2000. *The Emotions: A Philosophical Exploration*. Oxford: Clarendon.
- Goldman, A. 1991. Stephen P. Stich: The fragmentation of reason. *Philosophy and Phenomenological Research* 51: 189–93.
- Goldman, A. 1993. Epistemic folkways and scientific epistemology. *Philosophical Issues* 3: 271 – 85.
- Goldman, A. 1999. *Knowledge in a Social World*. Oxford: Clarendon Press.
- Goodman, N. 1955. *Fact, Fiction and Forecast*. Cambridge, Mass.: Harvard University Press.
- Grattan, L. M. and P. J. Eslinger. 1994. Long-term psychological consequences of childhood frontal lobe lesion in patient DT. *Brain and Cognition* 20: 185–95.
- Greco, J. 2003. Knowledge as credit for true belief. In DePaul, M. and L. Zagzebski (eds.) *Intellectual Virtue*, Oxford: Clarendon Press, 111–34.
- Greene, J. 2003. From neural “is” to moral “ought”: what are the moral implications of neuroscientific moral psychology? *Nature Reviews Neuroscience*, 4: 847–50.
- Greene, J. and J. Haidt. 2002. How (and where) does moral judgement work? *Trends in Cognitive Sciences* 6: 517–23.
- Greene, J., L. Nystrom, A. Engell, J. Darley and J. Cohen. 2004. The neural bases of cognitive conflict and control in moral judgement. *Neuron* 44: 389–400.
- Greene, J., R. Sommerville, L. Nystrom, J. Darley and J. Cohen. 2001. An fMRI investigation of emotional engagement in moral judgement. *Science* 293: 2105–8.
- Griffiths, P. 1997. *What Emotions Really Are: The Problem of Psychological Categories*. Chicago: University of Chicago Press.
- Grimm, S. 2008. Epistemic goals and epistemic values. *Philosophy and Phenomenological Research* 77: 725–44.
- Haack, S. 1997. ‘The Ethics of Belief’ Reconsidered. In M. Steup (ed.)

- Knowledge, Truth, and Duty*, Oxford: Oxford University Press, 21–33.
- Haack, S. 1998. *Manifesto of a Passionate Moderate*. Chicago: Chicago University Press.
- Hacker, P. M. S. 2006. Passing by the naturalistic turn: on Quine's cul-de-sac. *Philosophy* 81: 231–253.
- Hacker-Wright, J. 2009. Human nature, personhood and ethical naturalism. *Philosophy* 84: 413–27.
- Haidt, J. 2001. The emotional dog and its rational tail. *Psychological Review* 108: 814–34.
- Hanna, J. 2004. The scope and limits of scientific objectivity. *Philosophy of Science* 71: 339–61.
- Hanna, R. 2006. *Rationality and Logic*. Cambridge, Mass.: MIT Press.
- Hare, R. M. 1952. *The Language of Morals*. Oxford: Clarendon.
- Hare, R. M. 1957. Geach: Good and evil. *Analysis* 17: 103–11.
- Harman, G. 1982. Metaphysical realism and moral relativism. *The Journal of Philosophy* 79: 568–75.
- Harman, G. 1984. Positive versus negative undermining in belief revision. *Nous* 18: 39–49.
- Harrison, J. 1976. *Hume's Moral Epistemology*. Oxford: Clarendon Press.
- Heal, J. 1989. *Fact and Meaning: Quine and Wittgenstein on Philosophy of Language*. Oxford: Basil Blackwell.
- Hieronymi, P. 2008. Responsibility for believing. *Synthese* 161: 357–73.
- Hochschild, A. 1983. *The Managed Heart*. Berkeley: University of California Press.
- Hoffman, M. 2000. *Empathy and Moral Development*. Cambridge: Cambridge University Press.
- Hollis, M. and R. Sugden. 1993. Rationality in action. *Mind* 102: 1–35.
- Hooker, B. 2000. *Ideal Code, Real World*. Oxford: Clarendon Press.
- Hooker, B. 2007. Rule consequentialism and internal consistency: A reply to Card. *Utilitas* 19: 514–9.
- Hookway, C. 2002. Emotions and epistemic evaluation. In P. Carruthers (ed.) *The Cognitive Basis of Science*, Cambridge: Cambridge University Press, 251–62.
- Horgan, T. and M. Timmons (eds). 2006. *Metaethics after Moore*. Oxford: Clarendon Press.
- Horwich, P. 2006. The value of truth. *Nous* 40: 347–60.
- Hubin, D. C. 2001. The groundless normativity of instrumental rationality. *The Journal of Philosophy* 98: 445–68.
- Hume, D. 1777 / 1978. *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*, third edition. Edited by L. A. Selby-Bigge and P. H. Nidditch. Oxford: Clarendon Press.
- Hume, D. 1888 / 1978. *A Treatise of Human Nature*, second edition. Edited by L. A. Selby-Bigge and P. H. Nidditch. Oxford: Clarendon Press.
- Huss, B. 2009. Three challenges (and three replies) to the ethics of belief. *Synthese* 168: 249–71.
- Jackson, F. 1998. *From Metaphysics to Ethics*. Oxford: Clarendon Press.
- Jacquette, D. 1997. Psychologism the philosophical shibboleth. *Philosophy and Rhetoric* 30: 312–31.
- Jacquette, D. 2001. Psychologism revisited in logic, metaphysics and epistemology. *Metaphilosophy* 32: 261–78.

- James, W. 1884. What is an emotion? *Mind* 9: 188–205.
- Janvid, M. 2004. Epistemological naturalism and the normativity objection or from normativity to constitutivity. *Erkenntnis* 60: 35–49.
- Jennings, D, M. Lepper and L. Ross. 1981. Persistence of impressions of personal persuasiveness: Perseverance of erroneous self-assessments outside the debriefing paradigm. *Personality and Social Psychology Bulletin* 7: 257-63.
- Joyce, R. 2006. *The Evolution of Morality*. Cambridge, Mass.: MIT Press.
- Kalderon, M. E. 2005. *Moral Fictionalism*. Oxford: Clarendon Press.
- Kant, I. 1785 / 1996. *Metaphysics of Morals*, trans. M. Gregor. Cambridge: Cambridge University Press.
- Katz, J. 1990. *The Metaphysics of Meaning*. Cambridge, Mass.: MIT Press.
- Katzoff, C. 2001. Epistemic virtue and epistemic responsibility. *Dialectica* 55: 105–18.
- Kenny, A. 1963. *Action, Intention and the Will*. London: Routledge and Kegan Paul.
- Kim, J. 1988. What is “naturalized epistemology”? *Philosophical Perspectives* 2: 381–405.
- Kim, J. 2005. *Physicalism, Or Something Near Enough*. Princeton: Princeton University Press.
- Knobe, J. and S. Nichols (eds.). 2008. *Experimental Philosophy*. New York: Oxford University Press.
- Knorr Cetina, K. 1999. *Epistemic Cultures: How the Sciences Manufacture Knowledge*. Cambridge, Mass.: Harvard University Press.
- Kornblith, H. (ed.) 1994. *Naturalizing Epistemology*. 2<sup>nd</sup> edition. Cambridge, Mass.: Bradford Books.
- Kornblith, H. 1994. Naturalism: both metaphysical and epistemological. *Midwest Studies in Philosophy* 19: 39-52.
- Kornblith, H. 2002. *Knowledge and Its Place in Nature*. Oxford: Clarendon Press.
- Korsgaard, C. 1983. Two distinctions in goodness. *The Philosophical Review* 92: 169–95.
- Kripke, S. 1982. *Wittgenstein on Rules and Private Language*. Cambridge, Mass.: Harvard University Press.
- Kuhn, T. 1970. *The Structure of Scientific Revolutions*. 2<sup>nd</sup> edition. Chicago: University of Chicago Press.
- Kunda, Z. 1990. The case for motivated reasoning. *Psychological Bulletin* 108: 480–98.
- Kusch, M. 1995. *Psychologism*. London: Routledge.
- Kusch, M. 2006. *A Sceptical Guide to Meaning and Rules: Defending Kripke’s Wittgenstein*. Montreal & Kingston: McGill-Queen’s University Press.
- Lackey, J. 2007. Why we don’t deserve credit for everything we know. *Synthese* 158: 445–61.
- Lackey, J. 2009. Knowledge and credit. *Philosophical Studies* 142: 27–42.
- Ladyman, J. and D. Ross. 2007. *Every Thing Must Go*. Oxford: Clarendon Press.
- Laudan, L. 1990. Normative naturalism. *Philosophy of Science* 57: 44–59.
- Leslie, A. and S. Keeble. 1988. Do six-month-old infants perceive causality? *Cognition* 25: 265 – 88.
- Lewis, D. 1991. *Parts of Classes*. Oxford: Blackwell.
- Lowe, E. J. 1998. *The Possibility of Metaphysics*. Oxford: Clarendon Press.

- Lowe, E. J. 2006. Needs, facts, goodness, and truth. In S. Reader (ed.) *The Philosophy of Need*, Cambridge: Cambridge University Press, 161–173.
- Lynch, M. 2004. *True to Life*. Cambridge, Mass.: MIT Press.
- MacIntyre, A. 1959. Hume on “Is” and “Ought”. *The Philosophical Review* 68: 451–68.
- MacIntyre, A. 1999. *Dependent Rational Animals*. London: Duckworth.
- Mackie, J. L. 1977. *Ethics: Inventing Right and Wrong*. London: Penguin.
- Maddy, P. 1990. *Realism in Mathematics*. Oxford: Clarendon Press.
- Maddy, P. 1997. *Naturalism in Mathematics*. Oxford: Clarendon Press.
- Maddy, P. 2007. *Second Philosophy: A Naturalistic Method*. Oxford: Oxford University Press.
- Maffie, J. 1990. Naturalism and the normativity of epistemology. *Philosophical Studies* 59: 333–349.
- Margolis, J. 1997. Late forms of psychologism and anti-psychologism. *Philosophy and Rhetoric* 30: 291–311.
- Markosian, N. 2000. What are physical objects? *Philosophy and Phenomenological Research* 61: 375–95.
- Martin, C. B. and J. Heil. 1998. Rules and powers. *Nous* 32: 283–312.
- McDowell, J. 1994. *Mind and World*. Cambridge, Mass.: Harvard University Press.
- McGinn, C. 1983. *The Subjective View*. Oxford: Clarendon Press.
- McGinn, C. 1984. *Wittgenstein on Meaning*. Oxford: Basil Blackwell.
- Medawar, P. 1961. *The Art of the Soluble*. Harmondsworth: Pelican Books.
- Medawar, P. 1984. *The Limits of Science*. Oxford: Oxford University Press.
- Miller, A. and C. Wright (eds.). 2002. *Rule-Following and Meaning*. Chesham: Acumen.
- Miller, A. 2002. Introduction to Miller and Wright (eds.) 2002, 1–15.
- Miller, A. 2003. *An Introduction to Contemporary Metaethics*. Cambridge: Polity Press.
- Millgram, E. 2005. *Ethics Done Right*. Cambridge: Cambridge University Press.
- Millgram, E. 2005. Hume on ‘Is’ and ‘Ought’. In his 2005, 218–246.
- Millum, J. 2006. Natural goodness and natural evil. *Ratio* 19: 199–213.
- Moll, J., R. Oliveira-Souza, P. Eslinger, I. Bramati, J. Mourao-Miranda, P. Andreiuolo and L. Pessoa. 2002. The neural correlates of moral sensitivity: a functional magnetic resonance imaging investigation of basic and moral emotions. *The Journal of Neuroscience* 22: 2730–6.
- Moore, G. E. [1903] 1988. *Principia Ethica*. Amherst, NY: Prometheus.
- Moore, G. E. 1932. Is goodness a quality? *Proceedings of the Aristotelian Society Supplementary Volumes* 11: 116–31.
- Moran, R. 1988. Making your mind up: Self-interpretation and self-constitution. *Ratio* 1: 135–51.
- Moser, P. K. and D. Yandell. 2000. Farewell to physical naturalism. In Craig, W. L. and J. P. Moreland (eds.). 2000. *Naturalism: A Critical Analysis*. London: Routledge, 3–23.
- Nagel, T. 1970. *The Possibility of Altruism*. Oxford: Clarendon Press.
- Nagel, T. 1986. *The View from Nowhere*. Oxford: Oxford University Press.
- Nagel, T. 1997. *The Last Word*. Oxford: Oxford University Press.
- Nature Editorial. 2009. Humanity and Evolution, *Nature* 457: 763–4
- Naylor, M. 1985. Voluntary belief. *Philosophy and Phenomenological Research* 45: 427–36.

- Nisbett, R and L. Ross. 1980. *Human Inference: Strategies and Shortcomings of Social Judgement*. Prentice-Hall.
- Noë, A. 2004. *Action in Perception*. Cambridge, Mass.: MIT Press.
- Nottellmann, N. 2008. The present and future state of epistemic deontology. In Hendricks, V and D. Pritchard (eds.), *New Waves in Epistemology*, Basingstoke: Palgrave MacMillan, 75–105.
- O’Hagan, E. 2005. Belief, normativity and the constitution of agency. *Philosophical Explorations* 8: 39–52.
- Owen, D. 1999. *Hume’s Reason*. Oxford: Oxford University Press.
- Panksepp, J. 2003. Feeling the pain of social loss. *Science* 302: 237–9.
- Papineau, D. 1993. *Philosophical Naturalism*. Oxford: Blackwell.
- Papineau, D. 2003a. Normativity and judgement. In his *The Roots of Reason: Philosophical Essays on Rationality, Evolution, and Probability*. Oxford: Clarendon Press, 6–38.
- Papineau, D. 2003b. The evolution of knowledge. In his *The Roots of Reason: Philosophical Essays on Rationality, Evolution, and Probability*. Oxford: Clarendon Press, 39–82.
- Parfit, D. 2011. *On What Matters*. Oxford: Oxford University Press.
- Peirce, C. S. 1931. *Collected Papers, Vol. 1*. Cambridge, Mass.: Harvard University Press.
- Pelletier, F. J., R Elio and P. Hanson. 2008. Is logic all in our heads? From naturalism to psychologism. *Studia Logica* 88: 3–66.
- Perry, J. 1979. The problem of the essential indexical. *Nous* 13: 21.
- Perutz, M. 1998. *I Wish I’d Made You Angry Earlier: Essays on Science, Scientists and Humanity*. Oxford: Oxford University Press.
- Peterson, M. 2009. *An Introduction to Decision Theory*. Cambridge: Cambridge University Press.
- Pettit, P. 1992. The nature of naturalism. *Proceedings of the Aristotelian Society* 62: 245–66.
- Pettit, P. 1993. A definition of physicalism. *Analysis* 53: 212.
- Pigden, C. 1989. Logic and the autonomy of ethics. *Australasian Journal of Philosophy* 67: 127–51.
- Pigden, C. 1990. Geach on ‘good’. *The Philosophical Quarterly* 40: 129–54.
- Plantinga, A. 1993a. *Warrant: The Current Debate*. Oxford: Oxford University Press.
- Plantinga, A. 1993b. *Warrant and Proper Function*. Oxford: Oxford University Press.
- Plantinga, A. 2002. The evolutionary argument against naturalism. In Beilby ed., 2002, 1–14.
- Polanyi, M. 1962a. *Personal Knowledge*. Chicago: University of Chicago Press.
- Polanyi, M. 1962b. The republic of science: its political and economic theory. *Minerva* 1: 54–73.
- Popper, K. 1970. Normal science and its dangers. In Lakatos, I and A. Musgrave, eds, *Criticism and the Growth of Knowledge*, Cambridge: Cambridge University Press, 51–9.
- Price, B. H., K. R. Daffner, R. M. Stowe and M. M. Mesulam. 1990. The compormental learning disabilities of early frontal lobe damage. *Brain* 113: 1383 – 93.
- Pritchard, H. A. 1932. Duty and ignorance of fact. Reprinted in his *Moral Writings*, 2002, Oxford: Oxford University Press, 84–101.

- Prinz, J. 2004. *Gut Reactions: A Perceptual Theory of Emotion*. New York: Oxford University Press.
- Prinz, J. 2005. Are emotions feelings? *Journal of Consciousness Studies* 12: 9–25.
- Prinz, J. 2007. *The Emotional Construction of Morals*. Oxford: Oxford University Press.
- Prior, A. N. 1960. The autonomy of ethics. *Australasian Journal of Philosophy* 38: 199–206.
- Pritchard, D. 2007. Recent work on epistemic value. *American Philosophical Quarterly* 44: 85–110.
- Putnam, H. 1982. Why reason can't be naturalized. *Synthese* 52: 3–23.
- Putnam, H. 1992. *Renewing Philosophy*. Cambridge, Mass.: Harvard University Press.
- Quine, W. V. 1969a. Epistemology naturalized. In his *Ontological Relativity and Other Essays*. Columbia: Columbia University Press, 69–90.
- Quine, W. V. 1969b. Natural kinds. In his *Ontological Relativity and Other Essays*. Columbia: Columbia University Press, 114–138.
- Quine, W. V. 1974. *The Roots of Reference*. La Salle, Ill.: Open Court.
- Quine, W. V. 1980. Identity, ostension and hypostasis. In his *From a Logical Point of View 2<sup>nd</sup> edition*. Cambridge, Mass.: Harvard University Press, 65–79.
- Quine, W. V. 1981. *Theories and Things*. Cambridge, Mass.: Harvard University Press.
- Quine, W. V. 1986. Reply to Morton White. In L. E. Hahn and P. A. Schlipp (eds.) *The Philosophy of W. V. Quine*, La Salle, Illinois: Open Court, 663–665.
- Quine, W. V. 1992. *The Pursuit of Truth*. Cambridge, Mass.: Harvard University Press.
- Quine, W. V. 1995. Naturalism; or, living within one's means. *Dialectica* 49: 251–62.
- Quinn, P. 1994. The categorization of above and below spatial relation information by young infants. *Child Development* 65: 58–69.
- Radcliffe, E. S. 1999. Hume on the generation of motives: why beliefs alone never motivate. *Hume Studies* 25: 101–22.
- Railton, P. 1994. Truth, reason and the regulation of belief. *Philosophical Issues* 5: 71–93.
- Raley, Y. 2005. Ontological naturalism. *Pacific Philosophical Quarterly* 86: 284–94.
- Ramsey, W. 2002. Naturalism defended. In Beilby (ed.), 2002, 15–29.
- Ratcliffe, Matthew. 2007. *Rethinking Commonsense Psychology*. New York: Palgrave Macmillan.
- Rawls, J. 1999. *A Theory of Justice*. Revised edition. Oxford: Oxford University Press.
- Rea, M. C. 2002. *World Without Design: The Ontological Consequences of Naturalism*. Oxford: Oxford University Press.
- Reader, S and G. Brock. 2004. Needs, moral demands and moral theory. *Utilitas* 16: 251–266.
- Resnik, D. 1993. So scientific aims justify methodological rules. *Erkenntnis* 38: 223–32.
- Riggs, W. 2008. The value turn in epistemology. In Hendricks, V and D.



- Pritchard (eds.), *New Waves in Epistemology*, Basingstoke: Palgrave MacMillan, 300–23.
- Rind, M. and L. Tillinghast. 2008. What is an attributive adjective? *Philosophy* 83: 77–88.
- Rips, L. 1994. *The Psychology of Proof: Deductive Reasoning in Human Thinking*. Cambridge, Mass.: MIT press.
- Ritchie, J. 2008. *Understanding Naturalism*. Stocksfield: Acumen.
- Roland, J. 2007. Maddy and mathematics: naturalism or not. *British Journal for the Philosophy of Science* 58: 423–50.
- Rorty, R. 1995. Is truth a goal of inquiry? Davidson Vs. Wright. *The Philosophical Quarterly* 45: 281–300.
- Russell, B. 1935. *Religion and Science*. London: Thornton, Butterworth.
- Russell, B. 1962. *An Inquiry into Meaning and Truth*. Harmondsworth: Penguin Books.
- Sartre, J-P. 1962. *Sketch for a Theory of the Emotions*. London: Methuen.
- Sartwell, C. 1992. Why knowledge is merely true belief. *Journal of Philosophy* 89: 167–80.
- Saver, J. L. and A. R. Damasio. 1991. Preserved access and processing of social knowledge in a patient with acquired sociopathy due to ventromedial frontal damage. *Neuropsychologia* 29: 1241–9.
- Scanlon, T. M. 1998. *What We Owe to Each Other*. Cambridge, Mass.: Harvard University Press.
- Scarre, G. 2001. On caring about one's posthumous reputation. *American Philosophical Quarterly* 38: 209–19.
- Schickore, J. 2008. Doing science, writing science. *Philosophy of Science* 75: 323–43.
- Schroeder, M. 2007. *Slave of the Passions*. Oxford: Oxford University Press.
- Schueler, G. F. 1991. Pro-attitudes and direction of fit. *Mind* 100: 277–81.
- Schueler, G. F. 2009. The Humean theory of motivation rejected. *Philosophy and Phenomenological Research* 88: 103–22.
- Searle, J. 1964. How to derive “ought” from “is”. *Philosophical Review* 73: 44–58.
- Searle, J. 2010. *Making the Social World*. Oxford: Oxford University Press.
- Sellars, W. 1963. Philosophy and the scientific image of man. In his *Science, Perception and Reality*, London: Routledge & Kegan Paul, 1–40.
- Shafer-Landau. 2003. *Moral Realism*. Oxford: Clarendon Press.
- Shah, N. 2003. How truth governs belief. *The Philosophical Review* 112: 447–82.
- Short, T. L. 2002. Darwin's concept of final cause: Neither new nor trivial. *Biology and Philosophy* 17: 323–40.
- Siegel, H. 1996. Instrumental rationality and naturalized philosophy of science. *Philosophy of Science* 63: 116–124.
- Skorupski, J. 1989. *John Stuart Mill*. London: Routledge.
- Slote, M. 2003. Review: Natural goodness. *Mind* 112: 130–9.
- Smart, J. J. C. 1963. Materialism. *Journal of Philosophy* 60: 651–62.
- Smith, M. 1987. The Humean theory of motivation. *Mind* 96: 36–61.
- Smith, M. 1994. *The Moral Problem*. Oxford: Blackwell.
- Smith, M. 1995. Internal reasons. *Philosophy and Phenomenological Research* 55: 109–31.
- Smith, M. 2009. Desires, values, reasons and the dualism of practical

- rationality. *Ratio* 22: 98–125.
- Solomon, M. 2001. *Social Empiricism*. Cambridge, Mass.: MIT Press.
- Solomon, R. C. 2007. *True to Our Feelings*. New York: Oxford University Press.
- Sosa, E. 2007. *A Virtue Epistemology: Apt Belief and Reflective Knowledge, Volume I*. Oxford: Clarendon Press.
- Spelke, E. S. 1990. Principles of object perception. *Cognitive Science* 14: 29–56.
- Stanley, J. 2005. *Knowledge and Practical Interests*. Oxford: Clarendon Press.
- Steglich-Peterson, A. 2009. Weighing the aim of belief. *Philosophical Studies* 145: 395–405.
- Stenmark, M. 2001. *Scientism*. Aldershot: Ashgate.
- Stich, S. 1985. Could man be an irrational animal? Some notes on the epistemology of rationality. *Synthese* 64: 115–35.
- Stich, S. 1990. *The Fragmentation of Reason: Preface to a Pragmatic Theory of Cognitive Evaluation*. Cambridge, Mass.: MIT Press.
- Strawson, P. F. 1952. *Introduction to Logical Theory*. London: Methuen.
- Sturgeon, N. 2003. Moore on ethical naturalism. *Ethics* 113: 528–56.
- Suppe, F. 1998. The structure of a scientific paper. *Philosophy of Science* 65: 381–405.
- Swinburne, R. 1991. *The Existence of God (revised edition.)*. Oxford: Clarendon Press.
- Taylor, S. E. and J. D. Brown. 1988. Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin* 103: 193–210.
- Thagard, P. 2002. The passionate scientist: emotion in scientific cognition. In P. Carruthers (ed.) *The Cognitive Basis of Science*, Cambridge: Cambridge University Press, 235–50.
- Thompson, M. 2008. *Life and Action: Elementary Structures of Practice and Practical Thought*. Cambridge, Mass.: Harvard University Press.
- Thomson, J. J. 1997. The right and the good. *The Journal of Philosophy* 94: 273–98.
- Thomson, J. J. 2008. *Normativity*. Chicago: Open Court.
- Tolstoy, L. 1987. My confession. In O. Hanfling (ed.) *Life and Meaning: A Reader*, Oxford: Blackwell, 9–19.
- Turp, M. 2008. Naturalised Epistemology and the Normative. *Forum Philosophicum* 13: 335–47.
- Urmson, J. O. 1950. On grading. *Mind* 59: 145–169.
- van Fraassen, Bas C. 2002. *The Empirical Stance*. New Haven: Yale University Press.
- van Inwagen, P. 2006. What is naturalism? What is analytic philosophy. In Corradini, A, Galvan, S. and E. J. Lowe (eds.) *Analytic Philosophy Without Naturalism*, London: Routledge, 74–88.
- von Mises, R. 1951. *Positivism: A Study in Human Understanding*. Cambridge, Mass.: Harvard University Press.
- Weberman, D. 1996. Sartre, emotions and wallowing. *American Philosophical Quarterly* 33: 393–407.
- Weintraub, R. 1995. What was Hume's contribution to the problem of induction? *Philosophy and Phenomenological Research* 45: 460–70.
- White, M. 1986. Normative ethics, normative epistemology, and Quine's holism. In L. E. Hahn and P. A. Schlipp (eds.) *The Philosophy of W. V.*

- Quine, La Salle, Illinois: Open Court, 649–662.
- White, R. 2010. You just believe that because ... *Philosophical Perspectives* 24: 573–615.
- Whiting, D. 2010. Should I believe the truth? *Dialectica* 64: 213–24.
- Wiggins, D. 1998. *Needs, Values Truth*. 3<sup>rd</sup> edition. Oxford: Clarendon Press.
- Wiggins, D. 2006. *Ethics: Twelve Lectures on the Philosophy of Morality*. London: Penguin Books.
- Wiggins, D. 2009. What is the order among the varieties of goodness? *Philosophy* 84: 175–200.
- Wikforss, A. 2001. Semantic Normativity. *Philosophical Studies* 102: 203–226.
- Williams, B. 1981. Internal and External Reasons. In his *Moral Luck*, Cambridge: Cambridge University Press, 101–13.
- Williams, B. 1985. *Ethics and the Limits of Philosophy*. London: Fontana Press.
- Williams, B. 1989. Internal reasons and the obscurity of blame. In his *Making Sense of Humanity*, Cambridge: Cambridge University Press, 1995, 35–45.
- Williams, B. 2000. Philosophy as a humanistic discipline. *Philosophy* 75: 477–96.
- Williams, B. 2002. *Truth and Truthfulness*. Princeton and Oxford: Princeton University Press.
- Williamson, T. 2005. Armchair philosophy, metaphysical modality and armchair metaphysics. *Proceedings of the Aristotelian Society* 105: 1–23.
- Wilson, Catherine. 1995. *The Invisible World: Early Modern Philosophy and the Invention of the Microscope*. Princeton: Princeton University Press.
- Wilson, T. D. 2002. *Strangers to Ourselves*. Cambridge, Mass.: Belknap Press.
- Winch, P. 1983. Review: Facts and superfacts. *Philosophical Quarterly* 33: 398–404.
- Wittgenstein, L. 1958. *Philosophical Investigations*, trans. G. E. M. Anscombe. Oxford: Basil Blackwell.
- Wittgenstein, L. 1978. *Remarks on the Foundations of Mathematics*, trans. G. E. M. Anscombe. Oxford: Basil Blackwell.
- Wittgenstein, L. 1979. *Notebooks: 1914-1916*. Chicago: The University of Chicago Press.
- Wright, L. 1973. Functions. *Philosophical Review* 82: 139–68.
- Xu, F. 1997. From Lot's wife to a pillar of salt: evidence that *physical object* is a sortal concept. *Mind and Language* 12: 365–92.
- Younger, B and L. Cohen. 1983. Infant perception of correlations among attributes. *Child Development* 70: 291–303.
- Zagzebski, L. 1996. *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge*. Cambridge: Cambridge University Press.
- Zagzebski, L. 2003. The search for the source of epistemic good. *Metaphilosophy* 34: 12–28.
- Zimmerman, D. 2003. Why Richard Brandt does not need cognitive psychotherapy, and other glad news about idealized preference theories in meta-ethics. *The Journal of Value Inquiry* 37: 373–94.