

Realistic Fictionalism

Christopher Jay

UCL

(Submitted for the degree of PhD,
Summer 2011)

I, Christopher Jay, confirm that the work presented in this thesis is my own, except where indicated.

_____ (14/09/2011)

Abstract

Realistic Fictionalism, argues for two main claims: First, that there is no conceptual or logical incoherence in the idea of a fictionalist theory of some discourse which accommodates a form of realism about that discourse (a claim which has been made in passing by various people, but which has never been adequately explored and assessed); and Second, that just such a fictionalist theory promises to be the best theory of our ordinary moral commitments, judgements and deliberation.

In Part I, I explore the spirit of fictionalism and argue that thinking of fictionalism as closely tied to an analogy between its target discourse and *fiction* is liable to be misleading and is not mandatory. It emerges that the fictionalist's strategy requires just a semantic thesis (representationalism) and a thesis about the sort of 'acceptance' appropriate for some practice involving their target discourse (nondoxasticism). I offer a theory of what 'acceptance' is, which treats belief as a mode of acceptance and distinguishes the nondoxastic modes of acceptance from belief in a principled and independently plausible way. And I argue that the coherence of *realistic* fictionalism is preserved by the fact that a person (the realistic fictionalist) can perfectly coherently both believe *and* nondoxastically accept the same claims.

In Part II, I employ the theory of acceptance developed in Part I to propose a fictionalist model of how our ordinary moral commitments often are and generally ought to be. I then give an argument to the conclusion that, in respect of the relation between moral commitment and action guiding at least, it would be better if our moral commitments were to be nondoxastic. I then argue that *realistic* fictionalism offers a better way of explaining why we ought to have any moral commitments at all than a non-realist fictionalist theory could.

Contents

Abstract 3

Acknowledgements 6

Introduction: **Introduction** 9

1. Introduction 9

2. Some Key 'Isms' 14

3. The Character-Centred Fictionalist Model 24

4. The Thesis 31

Part I

Chapter One: **The Spirit of Fictionalism** 34

1. Introduction 34

2. Fiction, Truth and Acceptance 36

3. Different Types of Fictionalism 45

4. Fictionalism and Truth in Fiction 53

5. Conclusion 55

Chapter Two: **Fictionalism for Realists** 58

1. Introduction 58

2. Fictionalism 60

3. Daly's Trilemma 68

4. The Crowding Out Objection 70

5. No Crowding Out 72

6. Conclusion 83

Chapter Three: **The Varieties of Acceptance** 84

1. Introduction 84

2. Acceptance and Belief 86

3. A Theory of Acceptance 91

4. Horwich's Objection 103

5. Fictionalism and Pragmatic Norms 111

6. Not Just Practicalities 114

7. Conclusion 116

Part II

Chapter Four: **The Nondoxastic Character of Moral Commitment** 120

1. Introduction 120

2. The Character of Moral Commitment 124

3. Evidence 132

4. Excursus: A Literary Illustration 147

5. Conclusion 151

Chapter Five: The Good in the Nondoxastic	153
1. Introduction	153
2. Must Nondoxastic Commitments be Flimsy?	155
3. Fictionalisms: Hermeneutic, Revolutionary, and Evaluative	156
4. The Subjective Warrant Argument	162
5. Realistic Fictionalism and the Subjective Warrant Argument(s)	188
6. Conclusion	192
Chapter Six: Vindicating Realism	194
1. Introduction	194
2. Error Theory and the Challenge to Ordinary Commitments	198
3. How a Vindication Can Be the Wrong Sort of Vindication	202
4. The Wrong Sort of Vindication of Morality	208
5. An Adequate Vindication of Morality	214
6. Conclusion	229
Conclusion: Conclusion	231
References	236

Acknowledgments

The Philosophy department at UCL is a wonderful place to do research, and my time there writing this thesis has been filled with friendliness and lots of stimulating discussion, from which I have learned an immeasurable amount. This thesis would not be even as good as it is now were it not for the people and atmosphere there.

The debt I owe to Mark Kalderon's published work on fictionalism will be obvious from what I have written, but what I really want to thank him for is his support and advice in his capacity as supervisor throughout my writing it. He has been always pragmatic, frequently critical (but always constructively), and the source of a great many ideas and suggestions. And not only has it been exciting working on fictionalism with the person whose book, *Moral Fictionalism*, awakened me to both the topic and the possibility of a novel position within it worth exploring, it has been a valuable lesson in how to do philosophy to work at close quarters with a philosopher whose interests and expertise range so widely and so deeply.

I have also had the privilege of discussing drafts of some portions of the written work included here (or related work) with Marcus Giaquinto and with Jonathan Wolff. I want to thank them both for their encouragement and for their comments.

Much of my view, and of the way in which I develop it here, is the result of conversations I have had with others (sympathetic or sceptical) in the UCL department and within the University of London philosophy community, in a variety of contexts: I thank Luc Bovens, George Hull, Rory Madden, Mike Martin, Lucy O'Brien, Moises Vaca Paniagua, David Papineau, Ian Rumfitt and Barry Smith; and in particular I thank Craig French for his continuing interest in what I am doing and his eye for detail, and Lee Walters for his insights and for championing my cause. Both have been invaluable intellectual allies. I must also mention a Friday afternoon on which I had conversations with first Maarten Steenhagen and then Amber Corfield-Moore which finally (after several failed attempts to adequately formulate my view of the nondoxastic character of moral commitment) put me in a position to say what my positive view is. I have also benefitted from conversation, outside London, with Ralph Wedgwood. And, for as long as I have been thinking about philosophy, I have been challenged and prompted by Hayder Al-Mohammad, as much on the topics of this thesis as on anything.

It is indicative of the spirit of friendly cooperativeness I have so much enjoyed at UCL that the (now annual) summer Work In Progress (WIP) seminars, at which the graduate community discuss each others' work, are so well attended and useful. In a 2010 WIP we discussed a paper of mine on

the themes of this thesis (though of which only a few paragraphs have survived into the final work), and in 2011 the WIP discussed a penultimate draft of Chapter Four. I very much want to thank everyone who attended those discussions, and particularly Tom Avery, Henry Clarke, Amber Corfield-Moore, Craig French, Alex Geddes, Dave Holly, George Hull, Ben Martin, Jonny McIntosh, Ed Nettel, Tim Short, Maarten Steenhagen, Lee Walters and no doubt some others whose contributing I have failed in my duty to recall for their help.

Versions of Chapter Two were presented at graduate conferences in Manchester (Open Minds IV) and Leeds (CMM Graduate Metaphysics Conference) in the summer of 2009, and I am grateful to Phil Brown, Phil Letts, David Liggins, Jon Robson (my commentator at Leeds), Jason Turner, Robbie Williams, and several others at those institutions for their questions and feedback.

I am certain that I owe debts of gratitude for ideas and discussions which I have not managed to record here. If I have forgotten (because I failed to adequately record) the names of some who have helped me it is because I am slovenly and decadent, and not because I am ungrateful.

I have enjoyed the financial support of both the AHRC (who fully funded my doctoral research in 2008-10, having already funded my MPhil) and the Royal Institute of Philosophy (whose Bursary has supported the final year of my research). I am very grateful to those institutions, and to those at UCL who were instrumental in securing their support.

My family have always encouraged me, and for that I am enormously grateful. I hope the fact that I have finally managed to complete my education after all this time will be as much a source of pride to them as their support (both financial and moral) has been appreciated by me.

It is customary on occasions such as this for the author to say that they could not have written the work without the love and support of their partner. I consider it no great compliment to mine to saddle her with partial responsibility for the existence of this work, but without Amber Corfield-Moore's love and support its author would be incalculably less happy and less utterly delighted to be alive. I hope *that* truthful compliment does something approaching adequate justice to the gratitude I owe her, though I know that it does not.

And the man that has anything
bountifully laughable about
him, be sure there is more in
that man than you perhaps
think for.

*Ishmael, in Moby Dick
(Herman Melville)*

– Introduction –

Introduction

Where [‘The Myth of the Seven’] speaks of ‘*making* as if you believe that *S*’, I would now say ‘*being* as if you believe that *S*, but not really believing it except possibly per accidens’ (See Yablo, 2002a). Related to this, mathematical objects may exist for all I know. I do not rule it out that ‘ $2 + 3 = 5$ ’ is literally true in addition to being metaphorically true, making it a twice-true metaphor along the lines of ‘no man is an island’. I also do not rule it out that ‘ $2 + 3 = 5$ ’ is a maybe-metaphor, to be interpreted literally if so interpreted it is true, otherwise metaphorically. (Compare ‘Nixon had a stunted superego’, to use James Tappenden’s nice example.) I think the existence issue can be finessed still further, but the margin is too small to contain my proof of this.

Stephen Yablo, ‘The Myth of the Seven’¹

1. Introduction

Realistic fictionalism is realist in its commitments (at least according to some of the myriad senses of ‘realist’), and yet is still properly called a version of fictionalism. It is the purpose of the first part of *Realistic Fictionalism* to defend the – apparently counterintuitive – coherence of such a view. And it is realistic in the more everyday sense of affording a genuinely promising, non-illusory, way of thinking about how some things are. It is the purpose of the second part of *Realistic Fictionalism* to motivate the view as an account of moral thought.

In the first instance, the very coherence of a form of genuine fictionalism which is also a form of genuine realism ought to be interesting to anybody familiar with the way in which fictionalists are typically – indeed,

¹ Yablo, 2005. Yablo’s reference in the text is to his paper ‘Go Figure: A Path through Fictionalism’ (Yablo, 2002). The note to ‘The Myth of the Seven’ from which this passage is taken was added for publication in order to highlight some changes of mind on Yablo’s part since writing the bulk of ‘The Myth of the Seven’ in 1997.

pretty much *universally* – committed to the rejection of the sort of realism I have in mind. It is at least worth noting that there are other options not just in purely logical space, but in the space of reasonably well-motivated views about subjects of enduring interest to philosophers, in advance of knowing whether, for all their being reasonably well-motivated, they are going to turn out to be the *right* view to hold about some particular subject. We ought, at least, to have the options before us before we choose. But defending the coherence of realistic fictionalism involves more than just highlighting another option. It involves understating fictionalism itself in a more nuanced way than has tended to be done, getting beneath and beyond the well-worn analogies, metaphors and similes by means of which fictionalists and writers about fictionalism have tended to explicate their views and stating precisely what is essential to, and what is not essential to, those views.

Fictionalists typically occupy the non-realist ('antirealist', 'nominalist', 'irrealist'...) ground in debates about morality (Joyce, 2001), science (van Fraassen, 1980), mathematics (Field, 1980), possible worlds (Rosen, 1990), fictional characters (Brock, 2002), temporal parts (Kroon, 2001), abstracta (Yablo, 2000) or whatever the target of their theory is. There are two ways of coming to occupy such ground in a distinctively 'fictionalist' way: (i) to have some arguments against – or at least a strong presumption against positing – the existence of the types of objects or facts required for the straightforward truth of some target claims, literally construed, and to have some reason for preferring to treat the target claims as best construed literally, and thus to recommend (or diagnose) some way of accepting the target claims which is not *belief* (that is, a *nondoxastic* mode of acceptance); or (ii) to re-construe the target claims as involving some sort of 'according to...' operator in their semantics which is occluded on the surface. It is a delicate question to what extent proponents of the second strategy are doing the same thing as proponents of the first, but representatives of each (Field, from the first, and

Rosen and Brock from the second, for example) call themselves 'fictionalists'.²

Though it is common to find fictionalism *defined as* a non-realist position, it has occurred to some already that what is really distinctive about the fictionalist's view – what is essential to the most minimal conception of what counts as 'fictionalism' – is something in which, in principle, a realist could share. Thus it has been suggested more or less in passing by Brock, 2002, Kalderon, 2005a, Nolan, 2005 and Yablo (in the epigraph to this Introduction, and in Yablo, 2002: esp. fn1 and sec.12) that there might be room for a fictionalist account which doesn't rule out a realist metaphysics of the discourse in question (on the understanding of a 'realist' metaphysics I intend to operate with, which I discuss below). Serious consideration of the prospects for and limits of such a view, which I shall call *realistic fictionalism*, has been limited, however, by a widely held assumption that realistic fictionalism, even if coherent, is bound to be unmotivated. The root of this assumption is relatively easy to diagnose: every fictionalist currently in print is at least agnostic about the metaphysics of their target discourse, and determined that their agnosticism is the only appropriate response to the epistemic position in which we find ourselves, or is positively antirealist, nominalist or irrealist about their target discourse, and in both cases the arguments for fictionalism which are offered depend in some way or another and to some extent or other upon those metaphysical commitments. It is natural to assume then, without much further reflexion, that the only reasons anyone would have for being a fictionalist would be something to do with the rejection of realism.

² It is a moderately interesting fact (perhaps) about the sociology of taxonomy here that it tends to be philosophers working on problems more closely associated with what we might call 'pure' metaphysics – possible worlds, temporal parts, fictional characters – who opt for the second, semantic operator, approach, whilst those working on metaphysical aspects of what we might think of as *practices* – moral, mathematical or scientific investigation – tend to embrace the first, literalist, approach.

Stephen Yablo has said tantalising things about the possibilities for what seems to me to be realistic fictionalism, without actually developing them in a way committed to realism. Yablo's fictionalism is, in the end, of the more ontologically *agnostic* than realist type (see Yablo, 2000), like van Fraassen's, 1980 constructive empiricism, and I think it is Yablo's doubts about the extent to which the realism/non-realism issue *matters* that lead him to pay little attention to what the prospects are for full-blooded realistic fictionalism, so he seems to count amongst those who see realistic fictionalism as an ultimately unmotivated option (unmotivated because we have no good reason to defend a particularly robust version of realism at all, given the fictionalist option). Mark Kalderon has discussed the compatibility of moral realism with a form of moral fictionalism (in Kalderon, 2005a: ch4), and rejects the option not as unmotivated, but as too awkward a conjunction of views to be viable. That conclusion, though, is based upon a very particular understanding of what the moral fictionalist is to be committed to, an understanding which is very much optional.

So those who have glimpsed the very coherence of realistic fictionalism have, thus far, failed or refused to sympathetically assess its merits, and have sometimes, like those for whom the possibility of realistic fictionalism has never occurred, failed to see that there might be reasons to be a fictionalist which are not tied to the rejection of realism.

This thesis is about the possibility and merits of realistic fictionalism. Part I addresses various issues arising from the broadest, most general ways of thinking about fictionalism, and is intended to be of interest to those engaged in philosophical debates about *any* discourse for which fictionalism is an option. Chapter Three addresses a question which arises for fictionalists of any stripe, including realistic and non-realist fictionalists, namely what 'acceptance', of which fictionalists make a lot, actually is, and how belief is to be distinguished in a principles and independently plausible way from the

attitudes towards the target discourse that the fictionalist diagnoses or recommends. Chapters One and Two, however, are mainly aimed at correcting or forestalling some ideas about fictionalism which seem to be prevalent and which threaten to stand in the way of our having a clear view of the possibilities.

Part II is about metaethics, and addresses the assumption that realistic fictionalism is bound to lack motivation by aiming to provide just such a motivation, motivating and partially defending realistic fictionalism about our ordinary moral commitments, judgements and deliberation. At the very least, the reasons I discuss for treating our ordinary acceptance of moral claims and principles as often nondoxastic (in Chapter Four) and for thinking that our ordinary acceptance of moral claims and principles generally ought to be nondoxastic (in Chapter Five) have nothing to do with the rejection of realism, so those reasons are as available to the realist as to the non-realist. At most, if I am right that realistic fictionalism stands a better chance of vindicating the widespread and natural thought that there are some moral commitments we ought to have (which I argue in Chapter Six), then we have *some* reason to prefer realistic to non-realist fictionalism, if we are drawn to fictionalism at all (which, of course, I argue in Chapters Four and Five that we ought to be).

So Part I is intended to be of interest to those with no particular interest in metaethics, and Part II is *inter alia* a discussion of the nature of moral commitment which might be of interest to those with no axe to grind in any debates about fictionalism. Combined, though, Part II provides what remains lacking at the end of Part I, namely a *motivation* for a view such as the one Part I defends the coherence of; and Part I, in defending the coherence of such a view, addresses whatever doubts about the logical space for the sort of view proposed in Part II might otherwise arise.

2. Some Key 'Isms

Fictionalism, in the sense of 'fictionalism' I am interested in, is the conjunction of what I will call (*semantic*) *representationalism* and *nondoxasticism*.³ Nondoxasticism is discussed at length throughout this thesis. Representationalism is a semantic thesis about the content of the fictionalist's target claims: it says, very roughly, that their semantic content consists in — and is exhausted by — the stating or describing of putative facts. I assumed for a long time that I would write a chapter on semantic representationalism, but in fact I have nothing particularly original or interesting to say about it, and I have elected to restrict my chapters to discussions of ideas where I either have something to add or something to remind us of, of which we may have lost sight. But let me say a few things about representationalism and its role in this thesis now.

I take it that representationalists are opposed by *expressivists* of various sorts. Expressivism is the (class of) view(s) according to which the semantic content of some claims is *not* exhausted by their descriptive (representational) content, and that what completes the semantic content of those claims is something to do with the expression of noncognitive attitudes.⁴

The first serious appearance of expressivism on the semantic scene was with the emotivism typified by Ayer, [1936/46] and, more subtly, by Stevenson, [1944]. The emotivists used the language of 'evincing' attitudes (see, for example, Ayer, [1936/46]: p144) to explain that their view was not

³ My characterisation of fictionalism here closely follows Kalderon, 2005a; 2005c. He uses the label 'noncognitivism' for what I am calling 'nondoxasticism'. I think '*nondoxasticism*' better represents the centrality of *belief* (not other cognitive attitudes) to the thesis.

⁴ Representationalists oppose *any* semantic theory according to which semantic content is not exhausted by representational content. But their interesting opponents in the debates in which fictionalism seems to be a live option are those (expressivists) who share a (rough) thought about what sort of thing is required to complete semantic content.

that moral claims (which were their target) *describe* the attitudes of those whose claims they are, but rather that those attitudes are somehow *shown* by those claims.⁵ Of course it can be agreed on all sides that some moral claims (for example) show that a person adopts some attitude, and what representationalists deny is that that is a *semantically* relevant fact. Emotivists also placed great weight on the use or function of, for example, moral claims (to invoke attitudes in an audience, for example). Representationalists do not deny that the target claims *have* a use or function, but they deny that that use or function 'gets into the semantics'. And in refusing to read use or function into the semantic content of their target claims, representationalists avoid the 'pragmatic fallacy', the fallacy of moving straight from the *pragmatic* features of (the employment of) some claims to those claims' *semantic* features, diagnosed as at the root of emotivism by Dewey, 1945.

As if it weren't bad enough that emotivism was built on a fallacy (or, more charitably, *even* if emotivism were not built on a fallacy), emotivists and other early expressivists were soon seen to be subject to the notoriously brutal 'Frege-Geach Problem'. It became clear (and was pointed out by, for example, Geach, 1965) that if it really were the case that moral claims, for example, depend for their *content* upon the expression of some attitude, then occurrences of such claims in contexts in which no such attitude is expressed (such as in the antecedents of conditionals, for example) must be problematic: either such occurrences really involve different semantic entities to 'standard' occurrences (in which case it turns out that intuitively valid arguments involving embedded and non-embedded occurrences are invalid by equivocation); or, improbably, embedded moral claims must

⁵ Compare the notorious distinction between saying and showing in Wittgenstein, [1921]. I think, though this is a matter for elsewhere, that the early Wittgenstein is, in fact, best understood as a 'proto-emotivist'.

somehow be treated as expressing an attitude in the relevant way after all (but whose, and how?).

More recent expressivists, such as Blackburn, 1998 and Gibbard, 2003 have gone to impressive lengths to avoid the Frege-Geach Problem, by arguing that validity is preserved for arguments involving embedded and non-embedded claims on an understanding of validity that is broader than the notion assumed by proponents of the Frege-Geach Problem. Their theories (especially Gibbard's) are rich and complex, and deserving of careful treatment. Unfortunately, the best careful treatment I know, namely the one in Schroeder, 2008, makes a good case for thinking that expressivism for some particular discourse(s) is bound to be, at best, an outside contender. The problem for the sophisticated expressivist is that they address the familiar Frege-Geach Problem – and several other well-known problems for less sophisticated expressivism – by stressing the *similarity* between their target discourse and what they take to be purely representational discourse, arguing that the same semantic schemas can be applied to expressive and representational discourse, the differences being in the sorts of attitudes in terms of which those schemas are filled out (beliefs for the representationalist, noncognitive attitudes for the expressivist). But this idea, Schroeder argues with some ingenuity, means that sophisticated expressivists are bound to accept some very counterintuitive theories about concrete bits of *representational* discourse. The cost of carrying out the sophisticated expressivist project, if Schroeder is right, is very high indeed.

The point of this extremely quick and dirty run through (some highlights of) the debate in philosophical semantics between expressivists and their detractors has just been to indicate the reasons for which I am inclined against expressivism for most sorts of target claims. I have no new arguments against expressivism to offer – in my view, the naysayers of expressivism have very much the best of the debate as it currently stands,

and have tended, historically, to have the better of the debate. Whether Schroeder's conclusion that sophisticated expressivism seems ruinously revisionary of what we know of the semantics of representational discourse is the strongest conclusion to which we are entitled, or whether indeed there are even deeper problems with even sophisticated expressivism which mean that it fails to avoid even the problems widely acknowledged to put paid to less sophisticated expressivism whilst still counting as expressivism at all (see Kalderon, 2005a: ch2), is a matter for elsewhere.

That is one reason why I assume semantic representationalism (of which I have no particular favoured version in mind) in what follows. But another, even better, reason is that fictionalism is attractive to many of those who are attracted to it at all precisely because *inter alia* it rejects the expressivist account of the semantics of its target claims and offers *another* way of accommodating the noncognitive, or nondoxastic. As Kalderon, 2005a puts it, fictionalism is the prospect of 'noncognitivism without nonfactualism'. So, since my topic is fictionalism, it is quite appropriate that I should be concerned in what follows with only representational accounts of the semantics of our target domains.

Another 'ism which plays a large part in this thesis is *realism*. I want to say something about what I mean by it.

Even by the standards of philosophical terminology, 'realism' and its cognates are particularly widely, and thus confusingly, employed terms. By way of illustrating the bewildering range of meanings attached to the label, I'll merely note that it has been used for both a semantic thesis (by Dummett, for example, who says that 'in every case, we must regard a realistic view as consisting in a certain interpretation of statements in some class' (Dummett, [1982]: p230)) which (notwithstanding Dummett's particular views about the dependence of metaphysics on semantics) might admit of *any* sort of metaphysics, and a very particular sort of metaphysics which insists on the

existence of abstracta (this is often thought of as the ‘scholastic’ or medieval meaning of ‘realism’, but seems to be alive and well in Katz, 1998). For one very short survey of some uses of ‘realism’, see Neale, 2001: pp66-68. My use of ‘realism’ and ‘realistic’ accords with Neale’s second adumbrated sense, and seems to me to be common.

As I intend to employ the terminology, a realist is someone who thinks that a significant class of the claims of their target domain (so not just the negative existential claims concerning some contentious objects or facts, for example) are true, because there *are* such objects and/or facts as are required for their truth. So, a mathematical realist, in my sense, might be someone who thinks that since there are, for example, such things as sets, then there are true mathematical claims about them, and true claims which are best analysed as requiring sets for their truth. Such a realist is a realist about sets; and, given a theory of the ontology of mathematics according to which mathematical claims *are* reducible to set-theoretic claims, realism about sets is a species of mathematical realism.⁶

I say that those who oppose realists are *non*-realists. When Dummett introduced the ‘colourless’ term ‘anti-realism’ to stand for the opposite of realism (Dummett, [1964]: p145) he was no doubt right about its lack of colour. But since then – indeed, *because of* his introduction of that term in the context of his own discussion of realism, and the influence of that discussion – ‘anti-realism’ is all but impossible to disassociate from what we might call Dummett’s *semantic* interpretation of the debates between realists and their opponents. I shall say a little more about Dummett’s contribution, and its influence, below. But for now I simply want to point out that on my understanding of realism it is possible to oppose the realist by any of several means: by denying that any of the sorts of things required for the truth of the

⁶ Compare Boyd, 1990 for some doubts about the idea of ‘realism about’ some class of objects or facts.

target claims exist; by denying that we are in a position to assert (as the realist does) that they exist; or by denying that the truth of the target claims depends upon there being the particular sorts of things the realist is a realist about, for example. If we were eager to introduce yet more contentious terminology, we might say that the second way of resisting realism is to be 'agnostic', and then argue about how terms such as 'nominalist', 'anti-Platonist', 'irrealist', 'eliminativist', 'reductionist' and 'anti-realist' carve up the remaining options. I do not want to say anything much about that here, though. My point is just that the term 'non-realist' remains sufficiently colourless to convey the idea that the non-realist is opposed to realism without implying too much about how or why they oppose it.

It is tempting when thinking about what suffices for realism to appeal to ideas of mind-independence. Rosen, 1994 points out several difficulties with saying quite what such ideas amount to, though, and I am by no means confident that either *mind-independence* or *mind-independence* are the most general notions which might be of use. I shan't pursue that thought here, though, and if it seems indispensable to understanding talk of realism in my sense to think in terms of mind-independence then so be it. (I would, just as a final plea on this topic, though, urge the reader not to too hastily dismiss the option of replacing talk of *mind-independence* with talk of *judgement-independence* for the purposes of understanding one dimension along which realists and non-realists may be disagreeing.) However this is to be construed, realism in my usage throughout this thesis is a *metaphysical* thesis about what exists, how it is and whether it is as some target claims say that it is, on a more or less independent specification of what those target claims say.

My insistence on using 'realism' to stand for a metaphysical and *not* a semantic view (or range of views) is, of course, contentious. Again, I have nothing to say of my own to convince those for whom such a usage makes

little sense. I might appeal to Devitt, 1991 (and his defence of ‘Maxim 2’: ‘Distinguish the metaphysical (ontological) issue of realism from any semantic issue’ (1991: p3)) for support. At least it is nice to know I am not alone. But actually I am content not to go into battle over the whole issue of the relation of metaphysics to semantics (a battle which would require at least a whole thesis to make progress in), and I’ll settle for noting a few things about the Dummett-inspired tradition of thinking about realism which seem to suggest that that battle is strictly orthogonal, anyway.

When Timothy Williamson told off the philosophers working on realism for failing to hold themselves and each other to the highest standards of clarity and rigour (Williamson, 2006), he used Dummett’s construal of the debate between realism and its opponents as an example of the sort of precisification by which progress might be made, and bemoaned the fact (if it is a fact) that participants in the relevant debates have ‘tended to concentrate on the most programmatic issues, which they debated with no more clarity or conclusiveness than was to be found in the traditional metaphysical reasoning that Dummett intended to supersede’ (2006: p181). (Williamson’s point was not, in fact, that Dummett’s approach to metaphysics *via* semantics was the best approach, and indeed he rejects it himself, preferring, instead, to point out that at the very least semantics places some *restrictions* on the sorts of metaphysics we ought to find acceptable.)

But of course Dummett’s conception of realism was never intended to supersede *all* dispute about ‘realist metaphysics’ in the sense that has nothing much to do with taking sides over semantics. Explaining why he will not use the debate over what he is quite prepared to call ‘realism about universals’ as an example of the sort of *semantic* dispute he sees as at the root of many debates about various forms of realism in his original article, Dummett says ‘[i]t does not appear that the anti-realists in this case – the

nominalists – who denied the existence of universals and the referential character of general terms, were anti-realists in the sense of the characterisation I have now adopted: that they were necessarily committed to a different view of the kind of truth possessed by statements containing general terms (that is by all statements) from that of the realists’ (Dummett, 1964: p174). What Dummett did for the various debates about realism was, it seems to me, to provide a framework in which to discuss the difference between realists and their opponents in debates where it is agreed on both sides that some significant class of the claims at issue are true. Thus, in debates about the ontology of mathematics, certainly at the time Dummett became interested in these issues, it is often not doubted by either side that there are true mathematical claims. What divides Platonists (the realists in this debate) from their opponents (the intuitionists, on Dummett’s construal) is not the truth of mathematical claims, but what *explains* their truth. So in order to preserve the sense that there is a debate between *realists* and their opponents here, we had better construe realism as a thesis about what is required for the truth of some claims (or, in Dummett’s words, what sort of truth those claims have), a thesis with obviously semantic overtones.

Dummett’s conception of realism, then, is useful and important for characterising debates in which the two sides share a view about the truth-*value* of some claims, but differ over their explanation of that. As he later said, the ground of his strategy was that ‘since these metaphysical disagreements embodied divergent pictures of the reality to which the statements in question related, it seemed to me apparent that what underlay them were divergent pictures of the meaning of those statements’ (Dummett, [1992]: p465). Only if the participants in the debates in which Dummett was interested all thought that the statements in question were true would it be necessary (or desirable) to treat their disagreement as a disagreement about the *meaning* of those statements, for only if it were agreed on all sides that the

statements are true would it be the case that the natural place to look for their disagreement is in their account of the *bearers* of that truth; in so far as these metaphysical disagreements just embodied divergent pictures of the reality to which the statements in question related, it is an open question whether their disagreement is to be treated as a disagreement about meaning or about the truth values of those statements.

The point is that it would be unnecessarily and damagingly reductive to insist upon treating Dummett's semantic conception of realism as either the only properly intelligible conception or as the paradigm case. With the demise in popularity of verificationism and the arrival on the scene of theories (such as Field's, 1980 fictionalism) according to which it is entirely permissible to think that all the significant mathematical statements are *false*, the debate between 'realists' and their opponents in even the philosophy of mathematics no longer seems to require the semantic conception of realism to characterise it.

That is one reason why I take it to be open to me to appropriate 'realism' for the view I have in mind (or rather to fall in with what seems already to be a not-uncommon use of the term), the view that the target claims of some discourse are, in an important range of cases, true on some not excessively reductionist or eliminativist construal. But another, even better, reason is that my topic is fictionalism and the possibilities for a form of realism which is consistent with fictionalism. Fictionalists, at least in the sense I am interested in, are semantic *literalists*: they do *not* think that in order to preserve the truth of their target claims it is a good idea to give a semantics for those claims which eliminates reference to some problematic objects or facts, and they are not reductionists who think that talk of such object or facts is *really* referring to something less contentious, for they typically do not think that their target claims are true at all, precisely *because* they *do* refer (or putatively refer) to such problematic things. (This aspect of

the fictionalist's view is further explored throughout Part I). So if typical fictionalism is to count as non-realist, it must be non-realist on a construal of realism which doesn't count such semantic commitments as sufficient for realism. Since typical fictionalists are non-realists in virtue of their *metaphysical* commitments (their commitments about whether there *are* the things (putatively) referred to), then, realistic fictionalism will be fictionalism which is already 'realist' on the semantic conception of realism, but is also – and more interestingly – realist on the metaphysical conception: it will be the view that there *are* the things referred to, that they are the way they are said to be, and therefore that the target claims are true.

For all I've said about the way in which the conception of realism I am employing is not Dummett's semantic conception, there is a connection between representationalism and realism. On the conception of realism I have in mind, you need to be a representationalist to be a realist, though you don't need to be a realist to be a representationalist. I ought to say something briefly about some other 'isms, *quasi-realism*, and *creeping minimalism*.

Quasi-realism, as developed by Simon Blackburn (in, for example, Blackburn, 1993) aims to address the issues I have touched on in the representationalism/expressivism and realist/non-realist debates, but employing deflationary or minimalist notions of truth, reference etc. to say much the same things as representationalists and realists say without any of the hard metaphysical commitments to such things as, for example, moral truths or reference to moral facts or properties. The phenomenon of what James Dreier has called 'creeping minimalism' is the phenomenon of such quasi-realist theories appropriating *so* much of what the traditional realists say that it becomes hard to find anything over which quasi-realists and traditional realists disagree (Dreier, 2004). Far from being a welcome *rapprochement* between former enemies, this phenomenon only threatens to occlude a genuine debate, for quasi-realists are, by their own lights, only

quasi-realists because they reject much of what realists think. In order to rescue some hope of preserving a genuine disagreement, we must deny that the *quasi*-realists with their creeping minimalism about truth and reference are earning themselves what they claim. We must, that is, preserve notions of truth, reference, facts and the like which are too robust for the *quasi*-realist to swallow, and cast the debate between realists and non-realists (including *quasi*-realists) in *those* terms. We must also insist on a metaphysical conception of realism robust enough that the minimalists' talk of truth and reference is not sufficient for realism (see Asay, forthcoming: esp. sec.3 and 4).

In any case, fictionalism is naturally opposed to *quasi*-realism. Realistic fictionalists have no need for it, since they are full-blown realists; and non-realist fictionalists see *quasi*-realism as a competitor, not an ally. (Indeed, the feeling is mutual – see Blackburn, 2005.) So, since my topic is fictionalism and the coherence and merits of realistic fictionalism, I shall not be concerned with deflationary, minimalist or *quasi*-realist theories. My quarry is a form of realistic fictionalism far more interesting and contentious than any sort of *quasi*-realistic fictionalism with its desiccated notions of truth and reference, which would threaten to count as neither realism nor fictionalism on anything like the current usage.

3. The Character-Centred Fictionalist Model

So much for the myriad 'isms and epicycles of 'isms which we shall pass through in the course of this thesis but about which I have nothing more substantial to say. It might be helpful to have a glimpse of the fictionalist account of how our ordinary moral thought often is and generally ought to be which I shall be developing. The following sketch will hopefully give an idea of where we are going to be at the end (to give us hope of some

substantial results as we plough through Part I), but will, of course, advert to ideas which will only be fully worked out as we proceed.

The details of a fictionalist position can, of course, go any number of ways. The picture of moral commitment I propose is roughly this. It is a picture of *ordinary moral commitment and judgement*, meaning judgement (as opposed to passing moral thoughts, but (obviously) not limited to belief) about moral principles or facts as typically practiced in cases of moral deliberation and thought which goes on by people outside of the contexts of philosophical or other academic reflection on morality. It is a picture, if you like, of day to day morality. That is not to say that it is a picture of just our moral *intuitions*, even of our strongly held and most stable ones: I take it that moral thought can be something other (more?) than just intuition mongering without it being full blown philosophy of morality. On the picture I propose (which I'll call the 'Character-Centred' version of fictionalism) our moral commitments are often commitments to moral propositions, propositions which, if true, mean that there are moral facts, in some non-deflationary sense of 'fact'. But those commitments are not *beliefs* about those putative moral facts, so even if we think that there are no such moral facts, we have no reason to impugn our typical practice(s) of ordinary moral commitment with *error* on the basis of commitment to what is not really there, for it is only an error to *believe* what is false (compare: do we err if we accept that Sherlock Holmes was a detective for the purpose of literary discussion, even if we believe that he was not a detective, for nothing fictional could be?). The question now is what the nature of that commitment, if it is not belief, *is*.

The answer will, for our purposes now, have to be brief. It amounts to a view about what acceptance of a claim *is*, and how to differentiate various *modes* of acceptance. The thought is that to accept *P* is to treat *P* (i) as subject to some particular correctness condition, *C*, and (ii) as satisfying *C*. Believing *P* is a matter of treating *P* (i) as subject to a *truth-associated* correctness

condition (truth itself, or evidential warrant (or whatever)) and (ii) as being true, or evidentially warranted (or whatever). What is crucial here is the *internal* nature of the conditions mentioned by the schema and its instances: they are *not* (necessarily) the *actual* conditions which determine whether or not a belief (for example) *is* in fact correct. Nor is the perception of the relevant (internal) conditions being met necessarily anything to do with their *actually* being met. For all this characterisation of acceptance and its modes has to say, the actual correctness of (for example) beliefs may be determined by *external* conditions (causal, perhaps); that would just be for the correctness conditions *on* beliefs to be other than the conditions mentioned as being perceived to be relevant *in* believing. This will be important but it is not, I think, an *ad hoc* feature of the view: that there at least might be a distinction between what (for example) belief *is* and when it would be objectively *correct* is seemingly obvious, and the fact that my view mentions correctness conditions (in one case perceived, in the other actual) in characterising both of these distinct things should not tempt us to assume that that gap has *thereby* been narrowed.

What, then, of the mode of acceptance which *isn't* belief, which the Character-Centred version of fictionalism has as the relevant mode of acceptance in typical cases of ordinary moral commitment? Some fictionalists, and most critics of fictionalism, have thought that in spelling out the relevant nondoxastic mode of acceptance it is necessary to go beyond the *illustrative* function of the analogy with acceptance of fictional claims, and treat acceptance of (for example) moral claims *as* acceptance of the sort appropriate to the acceptance of fictional claims, or as something relevantly similar. I reject that thought.⁷ On my view, modes of acceptance are

⁷ To that extent, 'fictionalism' is an unfortunate name; but many pieces of philosophical terminology are misleading, and the sort of literalist nondoxasticism I have in mind is so close to versions of avowed fictionalism in the literature that it would be unnecessarily

differentiated by the different correctness conditions mentioned in the first clause of the definition of the mode (so, some truth-associated condition in the case of belief): all cases of acceptance amount to seeing whatever correctness condition is seen to apply as satisfied (the second clause), so the difference is in which particular correctness condition that is.

Plausibly (though nothing turns on this being right for either the Character-Centred view or the arguments I give), the relevant condition in the first clause of the specification of acceptance appropriate to fictional claims renders something like this: (fictionally) accepting P is to treat P (i) as correct iff some claim to the effect that P , or from which it can be inferred that P on the assumption that that claim is true, is to be found in a relevant text, and (ii) as satisfying *that* correctness condition. The mentioned correctness condition here is not truth-associated, for the only mention of truth is embedded in a mention of *assumed* truth, and we can assume true what is in fact false. So, if this characterisation of the mode of acceptance appropriate to fictions is right, that mode is not belief. Nor, though, is it what moral commitment is on the Character-Centred picture. But it is as well to linger briefly before saying what, according to that picture, moral commitment *is*. For the characterisation I gave of acceptance of fictions shows the difference between the doxastic mode of acceptance (belief) and nondoxastic ones. The difference is *not* that nondoxastic modes of acceptance necessarily amount to acceptance of some content which is not fully representational, for the nondoxastic characterisation of acceptance of fictional claims is acceptance *of* fully representational claims about such things as Sherlock Holmes. So nondoxasticism does not entail the denial of

peevish to demur from employing that name, so long as the name is treated as less of a guide to the content of the view than the explicit description which I give!

full semantic representationalism.⁸ The difference *is* that whilst it would make no sense for someone to say ‘I believe *P*, but I don’t think that *P* is true’, it would be perfectly intelligible (desirable, perhaps) for someone to say ‘I accept [in this particular way] that Sherlock Holmes was a detective, but I don’t think that it is true’. Whence that difference? From the difference of mentioned correctness conditions, which makes the former utterance (if sincere) ridiculous, but the latter (even if sincere) completely natural.

So, what is the character of ordinary moral commitment on the Character-Centred picture? On that picture, ordinary moral commitment to *P* typically involves treating *P* (i) as correct (in the context of moral deliberation) iff accepting *P* is required for being (or maximally approximating) a person with the sorts of character virtues one values, and (ii) as satisfying *that* standard of correctness. This picture is driven by the seemingly reasonable thought that a person can be moved to commit to some moral claim even though they do not know whether it is true, or even though they think that it is not true, because they insist on being *the sort of person who thinks*, for example, *that smacking children is wrong*. That insistence might plausibly just be an aspect of their conception of what the virtue of kindness requires, or is: a person could not be being properly *kind* if they were to countenance the smacking of children. And, crucially, the Character-Centred picture accommodates the reasonable thought that a person’s conception of the character virtues linked to moral commitment are truth-disinterested, that *kindness*, for example, is not good because of the relation it puts a kind person in to the truth, or to evidence, or to anything truth-

⁸ This is the crucial difference between fictionalism and standard breeds of noncognitivism (cf. Blackburn, 2005).

associated, as *rationality* is a good because, *inter alia*, it puts a person in a position to be sensitive to the right sort of evidence for their conclusions.⁹

Note again that these are what I called *internal* conditions: they are what it (often) *is* for a person to be morally committed, not what is it for their being so to be *correct*, for in seeing *P* as subject to *that* standard of correctness they *might* be making a mistake; and a person's *seeing P as* satisfying that correctness condition might similarly amount to a mistake, for it might be that accepting *P* is *not* in fact required for manifesting the virtue in question. (This is why the Character-Centred view of *moral commitment* does not presuppose or entail a virtue-theoretic account of any of the moral categories (good/bad; right/wrong; 'benevolent', 'cruel'...), for virtue theories are theories of what *is* good or wrong or benevolent or whatever.)

And note that on the Character-Centred view moral commitment is genuinely nondoxastic, for on it there is no incoherence – not even *prima facie* pragmatic incoherence – in a person saying 'Smacking children is [morally] wrong; but I don't believe in anything being objectively morally wrong', or anything which similarly manifests an opposition between what a person *believes to be true* and what they *treat as* being true in a truth-disinterested way. That opposition need not be, as it all too often has been, taken as

⁹ A reasonable worry about this model as a model of characteristically *moral* commitment is that there might be virtues of all sorts which matter to us. The worry is that we can take a virtue such as, e.g., being good at maths and then imagine a person who insists on being *the sort of person who thinks*, for example, *that seven is a prime number*, since thinking that seven is a prime number is required (they think) for manifesting the mathematical virtue. But we are surely not thereby imagining a person's *moral* commitments, so if Character-Centred commitment is typical of moral commitment, it is evidently not *sufficient* for it. (Nor of course is it *necessary*, on the plausible restriction of the Character-Centred view to a claim about how our moral commitment *often* are.) The reply to this worry ought to be this: The Character-Centred model is intended to model how our acceptance of moral claims often is, and it posits a distinctive mode of acceptance to play the required role; that mode of acceptance is distinctive in respect of its normative connection to perceived virtues and the perceived demands of those virtues; but it is distinctively *moral* only to the extent that the perceived demands of those virtues amount to acceptance of claims which are recognisably moral – and it is no part of the Character-Centred picture of moral *commitment* that there is any reason to abandon your favourite theory about what marks some *claims* out as distinctively moral.

tantamount to *expressivism*, which treats our commitments as expressive of some attitudes other than beliefs such that standard representationalism (or ‘descriptivism’) with respect to the *content* of our commitments fails; indeed, it is unclear whether such an opposition between what we believe to be true and what we treat as being true in a truth-disinterested way *can* reasonably be taken as tantamount to any such view of content, for it is not clear that nonrepresentational semantics is able to underwrite any notion of content fit for talk of treating some content *as* true. (Most contemporary expressivists are, of course, minimalists about truth, which, if itself a reasonable stance on truth, might assuage the worry I have gestured towards about the loss of truth-*aptness* concomitant with expressivism; but this is not the place to do any more than flag the issue.¹⁰) The fictionalist takes such an opposition as a consequence of accepting the (independently plausible) distinction between different ways of being committed to the *same* content – and since the way in which, on the Character-Centred view, moral claims are accepted is *opposed* to the strictures on acceptance associated with belief (strictures such that a person is incoherent to claim no commitment to the literal truth of what is believed), moral commitment, on the Character-Centred view, is nondoxastic.

¹⁰ Indeed, adopting some form of minimalism or deflationism about truth so as to entitle expressivists to some natural locutions is not new: Stevenson came to think (Stevenson, 1963b) that since it can be perfectly intelligible to say ‘that is true’ in response to some expression of a moral view, or to say of a person that ‘in my opinion their judgements are for the most part false’, it is necessary to adopt a ‘Ramsey-like’ theory of truth, namely, a deflationary one. (cf. Stevenson, 1944: pp169-71) (It ought to go without saying of course that expressivism is not essentially tied to any sort of deflationism or minimalism about truth. Gibbard, 2003 for one is distinctly ambivalent about the expressivist’s need for it, and see Nolan *et al*, 2005: p316. And of course there are truth minimalists (of some stripe) who are not expressivists, such as Wright (see, e.g., Wright, 1995).) There is an issue, going in the other direction from the issue flagged in the text to this footnote, about whether minimalism places any constraints on the formulation of expressivism – see the papers cited by Horwich, 1998: pp84-5, for the debate. But, as I said at the end of Section 2, minimalism is not my concern here.

What I have *not* said, of course, in characterising either what the possibilities for fictionalism are or what my preferred version of it amounts to, is that a fictionalist is bound to reject the idea that an important class of their target claims, complete with their literal ontological commitments, are *true*. The important class I have in mind are at least some of those target claims which, if true, would entail that there *are* moral facts or properties or whatever (not, that is, just the *negative* claims which even a denier of such things might endorse, such as ‘nothing is objectively cruel’).

Most fictionalists and critics of fictionalism think that a fictionalist’s attitude to realism is bound to range from hostility to indifference. Whilst fictionalism is commonly a response to some antecedent rejection of realism, a response to finding oneself driven to endorse an ‘error theory’ (in the sense of Mackie, 1977) about our typical employment of a particular discourse, the idea that our moral commitments generally ought to be as they are on the Character-Centred model, can be motivated without any such appeal to the rejection of realism.

4. The Thesis

The chapters which follow were written over a period of a little over two and a half years, and most started life as papers on particular ideas, problems or issues, all of which were intended all along to contribute to the larger project of exploring the prospects for realistic fictionalism, and motivating realistic fictionalism about our ordinary moral thought. In writing up this thesis, I decided not to rewrite the work completely, which would have allowed me to order the material more methodically and eliminate much repetition. I am painfully aware of the disadvantages of this decision, particularly the extent to which some points are repeated almost *ad nauseam* and the fact that the reader will not have a full picture of what I think fictionalism is until the end

of Part I. But I also have some faith that the presentation of the material which follows is apt to let the reader 'sneak up on' the main ideas in much the way I snuck on them in the course of development, and that the way things are done in this thesis represents not just what I think but why I think it and the interests which led me to form views at all. Of course the material here is often unrecognisable from the earlier drafts which are its ancestors, and I have deleted, moved and rewritten a great deal of material in putting the thesis together into its final form. Much of the repetition of key points which remains is the product of a deliberate (if tedious) policy of making clear at every relevant turn what bearing the ideas under discussion have to the broader themes of the thesis; hence, for example, my frequent encomiums to the possibility of realistic fictionalism which litter the discussions of all sorts of technical points throughout Part I.

I said, above, that I had assumed that I would write a chapter on semantic representationalism, but didn't in the end for fear of producing nothing more than a tedious survey of views, to which I have nothing much to add of relevance to my topic here. By contrast, I had not intended to write a chapter on the relation (if any) between *fictionalism* and fiction. But in the end I was driven to do so by the frequency with which conversations about fictionalism (and discussions in print) seem to appeal at some point or another to the assumption that a fictionalist about, for example, morality must be doing something like treating morality *as* a fiction. I take most of what I say in the chapter I ended up writing about that assumption to be rather obvious on reflection; but it seems sufficiently worth saying to be said first, before we embark upon further issues, so Chapter One now proceeds to explain 'The Spirit of Fictionalism'.

Part I

The Spirit of Fictionalism

‘The world is full of obvious things which nobody by any chance ever observes.’

Sherlock Holmes, in *The Hound of the Baskervilles*

1. Introduction

What sort of commitments does one need to have in order to earn the right to call oneself a ‘fictionalist’? And what, if anything, has fictionalism to do with fiction, or with the philosophy of fiction?¹

Fictionalists and writers about fictionalism sometimes appeal to analogies between how we do, or ought to, relate to some target discourse and how we relate to (typically narrative) fiction, in order to illustrate their proposal: just as we typically ‘accept’ but do not believe that Sherlock Holmes is or was a detective, we actually or ought to ‘accept’ but not believe the claims of some particular discourse (scientific, moral, mathematical etc.), and to prescind from believing the ontological commitments of that discourse just as we prescind from believing that there is or was a real (concrete, existing, person who is) Sherlock Holmes, even though the truth of many of the claims in the Sherlock Holmes stories, literally construed, requires that there be such a person. The analogy with fiction and our thought and talk about fiction is, obviously, a way of getting at the fictionalist’s core idea which makes the label ‘fictionalism’ very natural.

¹ It is worth noting that not all philosophers’ analogies between, say, mathematics and fiction are supposed to elucidate any sort of fictionalism. See, e.g., Wittgenstein’s (1978: IV.9) comment about the expansion of some term and whether the hero of some poem has a sister or not, and Wright’s (1980: pp168; 234-5) constructivist — not fictionalist — reading of this passage.

Similarly, the fictionalist move is sometimes characterised by way of an analogy with make-believe or pretence: just as children's play, for example, typically involves the pretence of some state of affairs, but not the literal *belief* (in those states of affairs) of the participants, we actually or ought to see the fictionalist's target discourse as involving some sort of pretence, withholding belief in the particular claims of that discourse (or, more accurately, the claims particular to that discourse). Of course it is not at all uncontroversial that make-believe and pretence are the same thing (see, for example, Sainsbury, 2009: ch1, esp. sec3), and if they are not then there are, in fact, even more options for the fictionalist who wants to illustrate their view by way of analogy – analogies with make-believe, *or* pretence, *or* fiction (if fiction isn't understood to itself be just a variety of one of these other phenomena, as it is by Walton, 1990, and the many who follow him, for example).

The fictionalists' analogies with fiction and pretence are certainly suggestive. But I want to explore precisely what it is about fiction that is so suggestive for fictionalist purposes. It turns out, I think, that focusing on the fiction analogy (and indeed focussing on any of the analogies I have mentioned) threatens to lead us astray from both the spirit and the letter of fictionalism, especially if we are insufficiently sensitive to the assumptions about fiction upon which those analogies are really based.

In the background is my idea that fictionalism is not as antithetical to *realism* as is usually supposed. In saying that fictionalism is not necessarily antithetical to realism, I mean something rather stronger than what is implicit in van Fraassen, 1980, whose version of fictionalism (if his 'constructive empiricism' is a version of fictionalism) amounts to *agnosticism* about the truth of (some of) the literally construed claims of science. My thought is that fictionalists can — and perhaps should — believe that their target claims are true rather than prescinding, in van

Fraassen's suggested way, from judgements about truth or falsity. One seeming reason for thinking that fictionalism and realism are necessarily incompatible, though (with respect to a particular discourse, anyway), is that any theory which invites comparisons between its target discourse and fiction or pretence must be committed to saying that the claims of the target discourse are, like the claims that constitute fictions, typically false. But I am not at all convinced that the *typical* falsity of fictional claims *need* be pertinent to a fictionalist. Rather than truth and falsity, distinctive norms governing acceptance seem to be the really interesting thing about fiction and our interaction with it. Realism is entirely at home with the sorts of fictionalist theories which build on *this* analogy between their target discourses and fictions.

Having more clearly understood what it is about fiction that is really distinctive, which might therefore be a more instructive point of analogy with some target discourse, the analogy itself starts to seem less precise and more contentious than might be hoped. So I want to suggest that whilst the analogy with fictions is suggestive, it is not the best way of conveying the distinctive content of fictionalism (at the most general level), at least once we move beyond relying on just our most everyday notion of fiction. We do better, I think, to characterise fictionalism about some discourse in terms of the distinctive conjunction of two theses, one semantic and one psychological. Once these theses are stated and understood, the analogy with fiction is merely illustrative at best (on *some* understandings of fiction) and misleading at worst.

2. Fiction, Truth and Acceptance

Sherlock Holmes could not be or have been a detective, for Sherlock Holmes does not and did not exist. That, at least, is what the seemingly

most natural theory of fiction — the one most of us who have not reflected deeply on the philosophy of fiction — tells us.² Nonetheless, utterances or inscriptions of ‘Sherlock Holmes was a detective’ can be *correct*.³ And the thought that Sherlock Holmes was a detective can quite properly play various roles for us, such as when we use the conceit to make a point (when, for example, we say things like ‘You don’t have to be Sherlock Holmes to work that out!’, or employ more vulgar colloquialisms in the vicinity). It is this relatively unusual disjoint between what is true and what it is, in some contexts at least, correct to think or say that many fictionalists focus on when drawing the analogy between their target discourse and fiction. So, for example, it might not be true that there are abstract objects of a particular sort (e.g. sets, or moral properties) but it could still be quite correct, on grounds of utility for example, to go on just as if there were.

Implicit in this analogy is the fictionalist’s commitment to the putatively fact-stating or descriptive nature of the target claims, which I take to be equivalent to their truth-aptness.⁴ Call this their commitment to full-blown *semantic representationalism*. What makes the fact that we can go on *as if* ‘Sherlock Holmes was a detective’ were true (even when we know that it is not) possible is that the claim is truth-*apt*. Were it not even truth-*apt* — if it were not the sort of thing that could truly describe the

² I discuss, below (Section 4), some of the relevant consequences of adopting other theories of fiction.

³ I say ‘utterances or inscriptions’ in order not to tread unnecessarily on any toes: perhaps it is alright to say ‘assertions of “Sherlock Holmes was a detective” can be correct’, or perhaps such an utterance or inscription could be correct only if it were a *quasi*-assertion of some sort. (See Kalderon, 2005a: ch3, for a discussion of various theories of quasi-assertion in this context.)

⁴ Of course this disagrees with various types of ‘minimalism’ about truth (e.g. Blackburn, 1984: ch6; Wright, 1992), which I shall not discuss here. Cf. Blackburn’s (2005) rejection of the idea that his quasi-realism is a form of fictionalism.

world — it would not be possible to so much as treat it *as* true.⁵ (It is perhaps this aspect of narrative fiction in its usual forms which marks the difference between it and some forms of poetry, for example, which we would not think of as if *true*, however engaged with them we are.) Similarly, fictionalists do not deny that their target claims are truth-apt, for they must preserve the possibility of treating those claims *as if* true (at least in some contexts, for some purposes). In this way, the analogy with fiction is well chosen.

Of course it is also implicitly built in to the analogy that the target claims are, typically, *false*. Fictions, after all, are commonly — and reasonably — contrasted with *facts*. But the matter is not so straightforward. Both with respect to the nature and the importance of fiction, truth — or rather falsity — does not seem to be as important as all that.

To see that falsity is not essential to the *nature* of fiction, just consider the case of a fiction which happens to say some things which are true. Very many of the things said in all sorts of novels are

⁵ Compare the possibility of going on as if ‘Hello’ were true. What is at stake here, I think, is the extent to which treating *as* true depends upon the possibility of *being* true. Clearly there are obstacles to truth which have no bearing on the possibility of treating as true — it is not possible (on one reading of ‘possible’, at least) that I am going to be able to fly tomorrow; nonetheless I can obviously treat the claim that I am going to be able to fly tomorrow *as* true. On the other hand, there are claims which cannot be true because they are not even *apt* for truth. Some of these are nonsense (‘Twas brillig, and the slithy toves did gyre and gimble in the wabe’). But the more interesting ones, so far as the point I want to mention about nonrepresentational semantics and treating as true is concerned, are meaningful in the way lines of poetry can convey meaning without depending for that on their (perhaps incomplete) descriptive content (as when, for example, adjectives are left hanging, unattached to objects for them to describe). I take it there is an interesting difference between novels, which are typically composed of claims which we can treat as true, and poems, which it often would make no sense to attempt to treat as *true*. That difference, I tentatively submit, might be to do with the representational semantics typical of novels and the nonrepresentational semantics of those poems (and those rarer novels, such as *Finnegan’s Wake*, perhaps) which defy treatment in respect of truth.

straightforwardly true.⁶ Perhaps most obviously, many of the claims made in ‘historical fiction’ are true – it is the fact that their plot is set in a recognisably accurate historical *milieu* that makes them *historical* fiction – but it doesn’t take much reflection to see that the phenomenon is pervasive across genres and types of fiction.

It is wildly implausible to think that there is some ‘critical mass’ of truth, more than which a novel cannot include if it is to count as fiction, so long as the work is presented as and functions as a fiction. Any line we proposed to draw would be bound to seem arbitrary. Similarly, it is hard to see what sort of principled way we could have of partitioning those claims in a fiction the truth of which is allowed for the fiction to still count as a full-blooded fiction from those claims the falsity of which is essential to the work of which they are a part counting as a fiction. There is a Kings Cross station in London; does that mean that the Harry Potter novels are not fiction after all? What if there really was a white whale which was the object of Ahab’s fascination? Would *Moby Dick* really not be a work of fiction then? The point here is not just that telling stories sometimes involves saying things which are true. It is that it doesn’t really matter *which* things are true and which are not: there is no obvious non-arbitrary line to be drawn between those ‘fictional’ claims which are allowed to be true and those which are not.

It might be stretching our intuitive notion of fiction to suggest that *every* claim a novel makes could be true and the work still be a work of

⁶ Recall that here, and throughout these sections until I turn explicitly to alternative theories of fiction in Section 4, I am exploring our more or less ordinary ideas about fiction. This seems reasonable, since the analogy with fiction is typically made on the basis of our more or less intuitive notion of fiction. And it is not typical of our normal everyday thinking about fiction to doubt whether, when a novelist says that Kings Cross is in London, they are saying something true. (Fictionalists all too rarely accompany their analogies with well worked-out discussions of the nature of fiction, and tend to rely on our more or less ordinary ideas. Perhaps an exception is Leng, 2010: esp. ch7, who offers a discussion of fictionalism, fiction and make-believe.)

fiction. (We might be tempted to call it an artwork of some other sort, if it were still presented and functioned in a non-standard way.) But so long as a work *is* a fiction (whatever that takes), any particular claims we accept as part of that work might be either true or false whilst remaining, plausibly, *fictional* claims, in virtue of their role in the fiction.

Perhaps we think that falsity is particularly necessary for fictional claims because the *significance* of fictions has to do with their not being mere reports of facts in the world. That would not, I think, be a good reason to assume that what is interesting about fictions for the purposes of analogy is that they are false. Even many reports of facts in the world are not *mere* reports. Spotlighting particular facts can be very important indeed: describing the psychological and social effects of slavery is not as mundane as writing up a telephone directory.⁷ That is to say, there is no reason to think that the moral significance, or the importance with respect to any of our purposes or interests, of a literally false description is generally any greater than that of a true description. In fact we might very well suppose the opposite. If the description of slavery (its conditions and consequences) occurred in a novel, would it matter to our reaction whether the person described happened to exist or not, or whether they really felt as they are described as feeling? Surely not, and if it did matter it would plausibly matter that they *did* exist or feel as described rather than that they *didn't*. In general, the assumption that the claims we are approaching when reading a novel, for example, are false does not seem to be playing much of a role in our engagement with a novel *as a fiction* at all. What matters is that the author is describing something to us, or telling us something, and that what we are reading either resonates or doesn't, either seems interesting or doesn't, is

⁷ The (moral) significance of this sort of description is nicely discussed in Wiggins, 2006 (esp. ch9).

pleasantly diverting or isn't, and so on. If someone were to stop us as we read and ask whether we thought the last claim we read is true, we would almost certainly say that *that just isn't the point*, even if we happen to think that as a matter of fact the last claim we read was not true.

I do not hope to have given anything like a comprehensive account of either the nature or importance of fiction. Indeed to do so in a properly philosophical way would, at this point, be inappropriate since my purpose for the time being is to explore the sort of *everyday* notion of fiction that is typically appealed to in the fictionalists' analogy. The point I hope to have brought out is that even though our ordinary notion of fiction, on which the fictionalist analogy typically rests, is typically committed to the falsity of fictional claims, that commitment is not constitutive of our notion of fiction in any particularly interesting way. We cannot, then, legitimately just assume that any analogue of fiction — such as the fictionalist's target discourse — necessarily involves falsity, for analogies are drawn between the particularly salient features of things, and the commitment to the falsity of particular fictional claims is not such a salient feature of our idea of fiction since it plays so little role in our thinking about the nature and importance of fiction.⁸

It is a mistake, then, to object to the coherence of realistic fictionalism on the grounds that such a view misappropriates the term 'fictionalism', *even if* (as I shall deny below) fictionalism is best thought of in terms of an analogy with fiction.

What, then, are the salient features which might properly ground the analogy? I think they are features of the norms governing acceptance

⁸ Stock, 2011 discusses some closely related themes, in particular the fact that even if fiction is to be understood as inviting the imagination of its audience, that does not mean that no fictional claims are intended to be believed, and are non-accidentally true. I take up some of Stock's themes in the following chapter, in which I discuss the relation between nondoxastic acceptance and belief.

of particular claims. What is distinctive about our acceptance of fictional claims is that we do not typically think that the correctness of our acceptance of them is subject to considerations of truth and falsity. Whether or not there was such a person as Sherlock Holmes, so whether or not we think it is true that Sherlock Holmes was a detective, it is correct for us to accept that Sherlock Holmes was a detective. And it is that distinctive feature which any informative analogy with fiction had better be playing on. It is what makes sense of the fictionalists' emphasis on the pragmatic virtues of accepting what they typically believe to be false, literally construed (or, in van Fraassen's case, for example, what might be either true or false for all the working scientist ought to care — by being officially uninterested in truth or falsity, on literal construals, van Fraassen is closer to the best sort of analogy with fiction, though he does not embrace the term 'fictionalism').

In Chapter Three, I shall say much more about acceptance norms. Suffice it for now to point out that when we read about Sherlock Holmes and his detecting exploits, it is *appropriate* for us to go on in certain ways: responding in certain ways to questions like 'What did Sherlock Holmes fill his days with?', for example. Notice that this is not true of the sort of 'acceptance' involved in *assumption*, for example. If we are assuming something for the sake of argument, we accept it only conditionally — we must, after all, discharge our assumptions when doing formal derivations. The 'loose' acceptance norms involved in fiction — 'loose' because they are not concerned with truth on literal construals — are not like that. When we accept that Sherlock Holmes is a detective, we do not just *conditionally* accept it, we accept it in the only way it is appropriate *ever* to accept it (*qua* fictional claim, at least).⁹ We do not, that is, just

⁹ The idea here is very close to that of 'full acceptance' found in Harman, 1986 and Kalderon, 2005a.

assume that it is true in advance of finding out whether it is. Indeed, we do not assume *that it is true* at all. And finding out whether it is true is not going to make our acceptance of it *qua* fiction any more or less appropriate.

As I will suggest below, whether the best *philosophical* theory of fiction fits the everyday notion of fiction I have been concentrating on in order to bring out this feature is an open question, at least for all that's been said so far. The most that any fictionalist can hope for from the analogy with fiction, I think, is the availability of a case (fiction) for which the 'looser' acceptance norms I have mentioned are less controversial than for the fictionalist's target discourse, which is typically a discourse to which it is typically assumed the standard, truth- or evidence-sensitive belief norms apply. (To this extent, to return to the issue of truth and falsehood, it would not matter if it *were* constitutive of the idea of a fiction that its claims must be false, for it would not follow that the *norms* in which the fictionalist is interested say anything about truth or falsehood: the norms are norms of acceptance, and they say that acceptance can be appropriate *whether or not* some claim is literally true; they do not, on any plausible interpretation, say that acceptance is appropriate *only if the claim is false!*)

The norms governing acceptance of fictional claims are distinctive, but I do not want to say — and I don't think any fictionalist should say — that fiction constitutes the only case of such norms. Similar norms are at work in treating something as pretence, or make-believe.¹⁰ Indeed, some fictionalists and writers about fictionalism (e.g. Kalderon, 2005a) draw an analogy between the target discourse and make-believe rather than

¹⁰ For the purposes of my brief discussion here I treat pretence and make-believe interchangeably. As I noted above, there are doubts about whether they are the same thing. But for all I want to say here, whatever differences there may be will not seriously threaten my point.

between the target discourse and fiction as it occurs in novels. But similar considerations apply to the analogy with pretence as apply to the analogy with fiction. Pretence in general does not *essentially* involve falsehood, though pretences are *typically* false. But to the extent that pretences are typically false, their *importance* or *purpose* is not essentially tied up with their being false.

To see that pretence or make-believe does not necessarily involve falsehood, think about variants of the true fictional claims I mentioned above. I might, for example, play the role of a confused PhD student in a play, written by a playwright with no knowledge of my predicament. The fact that I *am* a confused PhD student does nothing to undermine the thought that what I am doing on stage is *playing a part*: the pretence in which I am engaged is not undermined *qua* pretence by the fact that the pretence involves claims which are true. Or we might understand the make-believe model in terms of game-playing of the sort children engage in. Suppose that I am a small child playing a game of 'mummies and daddies', and that on this particular occasion I am playing at being the child. The fact that I really am a small child does no violence at all to the role that make-believe is playing in my participation in the game. (Indeed it does not even seem that I need to pretend to be a *different* child: what if I'm playing a game in which *I* am going to visit the pretend grown-ups for tea?)

As in the case of fiction, the significant aspect of the pretence involved in these sorts of situations seems to be that the norms governing acceptance of the claims being made (literally or tacitly) are distinctively loose. The appropriate mode of acceptance in these cases is not belief, but some mode of acceptance with a looser connection to what's true in the world or with no such connection at all. Whatever is true of the actor on stage, I ought to accept but ought not to believe what is said about him on

stage *qua* lines in the play.¹¹ Whatever is true of the children playing a game, those children (and anyone else indulgently going along with the make-believe) ought to accept but not really believe that things are such-and-such a way. Distinctively non-truth-based acceptance norms seem to be doing all the work in these cases. It is these norms which really distinguish pretence from normal assertion or belief. So it is these norms which seem to be the best reason for going out of one's way to draw an analogy with pretence.

Again, then, it is a mistake to object to the coherence of realistic fictionalism on the grounds that such a view misappropriates the term 'fictionalism', *even if* fictionalism is best thought of in terms of an analogy with pretence or make-believe.

3. Different Types of Fictionalism

Fictionalism is a broad church (see Eklund, 2009, for a fuller taxonomy), divided by a common insight. That insight is that the ontological commitments of some useful or otherwise virtuous discourse need not saddle those who would like to avail themselves of the benefits of that discourse with those ontological commitments themselves.¹²

¹¹ That is, if I have good reason to literally believe something that is said about an actor on stage, it had better be that I have some *independent* reason for believing it.

¹² As Daly, 2008, points out, no fictionalist ought to think that fictionalism – *any* version of fictionalism, in fact – avoids saddling *someone* (the fictionalist themselves, or the relevant practitioner with a discourse; a distinction Daly fails to make) with *some* ontological commitments (or some commitments to abstracta, in particular). But it is really not clear whether or why anyone would hold the view that fictionalism is a panacea for ontological commitments (or commitment to abstracta) *tout court*; all the fictionalist is typically concerned with is some particularly *problematic class* of ontological commitments (to numbers or sets, or postulated unobservables, or moral properties, etc.). And needless to say, *which* types of (commitment to) things are problematic depends on all sorts of scruples which vary from one to another. So *global* fictionalism would, if it were ever held, be a very surprising doctrine indeed, regardless of its eventual coherence.

Talk of ‘ontological commitment’ is usually restricted to a concern for what (sorts of) things *sentences* or *claims* are committed to the existence of. But ontological commitment is a species of *commitment*, and commitments are incurred by various sorts of things. As Stephen Neale notes (in passing, not stopping to consider its implications), ‘[w]e might think of the ontology of a sentence as those things that must exist for the sentence to be true; and we might think of the things in “our ontology” as the ontology of all the sentences we hold true’ (2001: p33). Perhaps Neale intended to be referring (by ‘our ontology’) just to the ontological commitments of a set of sentences (the ones we just happen to hold true), and to treat ‘our’ ontological commitments just as a subset of the ontological commitments of sentences as a whole. But I think it is clear, even from the way Neale actually puts the point, that *our* ontological commitments are a *function* of the ontological commitments of some subset of all the sentences, and that it is not the case that they *just are* that subset. For our ontology is determined by the sentences we *hold true* – our holding *S* true is as much a necessary condition on *our* being ontologically committed to whatever *S* is ontologically committed to as the content of *S* (it’s ontological commitments) is.¹³ It makes sense, then, to be clear that our ontological commitments depend upon the ontological commitments of the claims we hold true, but that they also depend upon *our holding them true*, in such a way that *we* can avoid the

¹³ Note also the following from Quine’s classic, ‘On What There Is’: ‘We can very easily involve ourselves in ontological commitments, by saying, e.g., that *there is something* (bound variable) which red houses and sunsets have in common; or that *there is something* which is a prime number between 1000 and 1010. But this is, essentially, the *only* way we can involve ourselves in ontological commitments: by our use of bound variables.’ (1948: p31) All I mean to draw attention to here is Quine’s speaking of involving *ourselves* in ontological commitments, and *our use of* bound variables – the point being just that the issue of what ontological commitments are incurred by particular *claims* is likely to be of interest *because* we are interesting in *our* commitments, in what acceptance of those claims means for *our* ontology.

ontological commitments of claims we accept so long as we accept them in ways that don't involve holding them true. This, of course, is the fictionalist's characteristic move with respect to ontological commitment: the literally construed ontological commitments of mathematical discourse, for example, don't commit mathematicians to the sort of error involved in thinking there *are* ontologically suspect mathematical objects.

There is, I think, good reason to agree with Sainsbury (2009: ch8) that in order to be deserving of a title all to itself, fictionalism had better not just be semantic reductionism or eliminativism, nor any form of semantic replacement theory, for such theories, which attempt to alleviate the seemingly problematic ontological commitments of some discourse by arguing that those are not in fact its real commitments, are more properly seen as seeking to *remove* the problem of our apparent need to be committed to some problematic types of things than to solve it. (Sainsbury helpfully notes, for example, that it makes little sense to call Berkeley a fictionalist, precisely because on his theory 'material' objects *are* just mental entities.) Nonetheless, the ways in which avowed fictionalists have sought to avail themselves of the fictionalist insight do not form a homogenous class. There are fictionalists, for example, who seem to be offering what looks very much like a semantic replacement but who perhaps still count as fictionalists because on their theory the only addition to the semantics for the target discourse is some sort of *according to...* operator, where the space is filled in by a theory or body of doctrine which *is* committed to some problematic type of things (see, for example, Rosen's, 1990, modal fictionalism, which treats claims about possible worlds as claims about what Lewis' modal realism says about possible worlds¹⁴). Much discussion has been generated by these sorts of

¹⁴ Rosen is admirably clear in his presentation of the basic principles underlying his strategy. He points out that 'as we ordinarily understand [claims within an *according to*

fictionalism, to the extent that it is now common to come across people reasonably familiar with the literature who think that fictionalism just is the view that some types of claims are true, if they are true at all, only 'according to the fiction' of some theory or body of doctrine.

Other, arguably better, ways of availing oneself of the fictionalist insight are available, though. Joyce, 2005, presents some good reasons for thinking that the 'operator' approach is inadequate. For one thing, it is hard to see why a non-realist fictionalist who is engaged themselves in their target practice would be inclined to withdraw their utterance if pressed on the real metaphysical status of their discourse: a non-realist who has said, as part of the practice of mathematical discussion, for example, that there is a prime number between five and ten would have no reason to say, when pressed on the existence of numbers, that what they said was actually false because there no numbers if what they said was, in fact, not that there is a prime number between five and ten but just that *according to the mathematical fiction* there is a prime number between five and ten. Yet fictionalists – non-realist ones, at least – tend to give the impression of being prepared to affect just such a withdrawal of

the story of... prefix], quantification within the scope of a story prefix is not existentially committing. You can believe 'According to the fiction F , $\exists xPx$ ' without believing ' $\exists xPx$ '; for as a rule, the former does not entail the latter.' He then suggests applying this lesson to talk of possible worlds, noting that 'the prefix will sometimes be *silent*; so the fictionalist will often sound just like the modal realist. Yet the fictionalist's claims about possible worlds will always be elliptical for claims about the content of a story; and the ellipses can always be expanded. In this way, the fictionalist hopes to earn honest title to the language of possible worlds – that is, title to talk as if there were such things – while retaining a sensible one-World ontology' (1990: p331, p132).

(Note that Rosen, like Daly, 2008 and Nolan *et al*, 2005, treats the *person who is engaging in the target discourse* as the 'fictionalist'. There is good reason, though, to think of fictionalism as a theory about the target discourse, and a 'fictionalist' as someone who believes that theory, regardless of whether or not they are themselves discourse practitioners – you do not have to be a mathematician to believe a theory about the existential commitments of mathematical claims (that is a *philosophical* question, and not all mathematicians are philosophers of maths). That is just something to bear in mind, though, and does not affect the point being made in invoking Rosen here. See Chapter Two.)

what is said in the target practice if the context shifts to one in which *truth* is germane, an impression they would have no reason to give on the operator approach since what has been said, on that approach, is perfectly true: *according to the mathematical fiction*, there is a prime number between five and ten.

For another thing, a fictionalist who thinks that we are or ought to be expressing claims about what is the case *according to the fiction* of something or other is not giving a particularly charitable account of what our commitments are, for on the operator theory simple and intuitive inferences threaten to turn out not to be valid. Suppose, for example, that one is a fictionalist about colour. (Joyce, 2005: p292f; Vision, 1994.) Now take a simple argument such as:

- (P1) Fresh grass is green;
- (P2) My lawn is made of fresh grass;
- (C) Therefore, my lawn is green.

The operator fictionalist faces a dilemma: either they just treat (P1) as actually expressing the claim that *according to the colour fiction* fresh grass is green, in which case the argument is invalid; or they preserve validity by insisting that the *according to the colour fiction* operator attaches to what is actually expressed by (P2) of the argument too, at the cost of having to say some very implausible things about what the content of the colour fiction is (why would it be part of the colour fiction that my lawn is made of fresh grass?).¹⁵ Joyce entertains the thought that it *might* be part of the colour fiction that all sorts of non-colour-related things are true, in the

¹⁵ Here, of course, is an analogue of the Frege-Geach problem for expressivists; so if fictionalism is supposed to be an advance on expressivism in part because of its avoiding the Frege-Geach problem, then *a fortiori* the fictionalist had better not adopt the operator approach.

same way as it might be true in any story that $2 + 2 = 4$, whether or not that is ever explicitly stated in the story, but notes quite reasonably that we are committed to facts about the composition of our lawns in all sorts of contexts, and that *any* of these commitments could be used in an argument such as the one above, so that it would always be possible to construct an invalid argument in the way described unless *all* of our (putative) commitments to facts about the composition of our lawns were really commitments to facts about the composition of our lawns *according to the colour fiction*, which is wildly implausible. So, Joyce thinks, better to formulate our fictionalism in some other way, a way which doesn't confuse *reporting on* the fiction with *participating in it*. I agree.

The sort of fictionalism I am interested in discussing and defending in Part II is not a semantic replacement strategy at all: it is committed to the *semantics* of a target discourse being just as any realist would like them to be (see, for example, Field, 1980; 1989 (esp. ch1), for a similarly 'literalist' sort of fictionalism.) The trick is turned, on this sort of fictionalist strategy, by treating the *acceptance* of the target claims as something other than belief. The semantic replacement strategist, of either a fictionalist or non-fictionalist sort, requires no such commitment about the nature of acceptance for, since they have cleansed the semantics of their target discourse of literal commitment to problematic types of things, believing their target claims is not bound to incur commitment to those types of things (see Rosen, quoted in n14, above). The literalist, if they are to avoid our being committed to problematic types of things in our acceptance of their target claims, however, must say instead that the very nature of that acceptance is such as to fail to transmit the ontological commitments of the target discourse, literally construed, to those who accept claims of that discourse. On this model of fictionalism, inspiration is still typically drawn from our engagement with fiction; but here the

emphasis is not on the thought, for example, that some claim is true *according to the Sherlock Holmes stories*, but rather on the thought that our acceptance of claims about Sherlock Holmes does not typically amount to belief, for we know that claims such as ‘Sherlock Holmes was a detective’ are bound to be false, and we do not believe what we know to be false. The fictionalist’s point is that just as we do not forswear accepting that Sherlock Holmes was a detective in *some* way just because we know that it is false, we might very well be accepting mathematical, or (some) scientific, or moral claims (etc.) regardless of whether we are committed to their truth, or to there even being the types of things required for their truth.

Another division amongst fictionalists is between ‘hermeneutic’ and ‘revolutionary’ fictionalists. As I shall explain in Chapter Five (Section 2), this classification, as commonly understood, is not exhaustive. Usually, hermeneutic fictionalists are taken to be proposing a theory of what our practice with some discourse is already like, to be merely *describing* what our moral commitments, for example, are like. They are usually contrasted with revolutionary fictionalists, who are proposing a theory of how we *ought to change* in our practice with some discourse, because our current practice is unsatisfactory.

It is not helpful, I think, to represent the distinction between hermeneutic and revolutionary fictionalists as simply a reflection of the distinction between description and assessment. It is not that I am sceptical about the radical distinction between description and assessment as activities; I am not. It is rather that what differentiates hermeneutic and revolutionary fictionalists is not that revolutionary fictionalists unlike hermeneutic fictionalists are interested in assessing our current and possible practices: the hermeneutic project is concerned with assessing its proposed analysis of our practice in the interests of

charity, after all – a theory about what our practice is like which represents it as hopelessly deviant where other hermeneutic theories would not is at least *prima facie* less plausible a representation of our settled and well-established practices, after all. The difference is rather that revolutionary fictionalists, unlike hermeneutic fictionalists, are interested in recommendation rather than description.

In the remainder of Part I, I shall not be very interested in this distinction, and I shall tend to speak of the fictionalist as someone with a theory about how our practice with a discourse ‘is or ought to be’, in order to allow my discussion to range over both hermeneutic and revolutionary fictionalism. In Part II, this distinction, and the distinction between those types of fictionalist theory and another type of fictionalist theory I call ‘evaluative’ fictionalism, becomes more important, and the fictionalism I motivate and defend in Part II is a theory about how our practice of ordinary moral deliberation ‘often is and generally ought to be’. The fact that I proceed in Part I largely without differentiating hermeneutic from revolutionary fictionalism is not meant to imply that I do not think that distinct issues arise for each (some of these issues arise in Part II); it is just that the ideas about the possibilities for fictionalism that I want to pursue in Part I are, I think, sufficiently general to apply to both, and frequent digressions on the specific issues arising from applying those ideas to hermeneutic or revolutionary fictionalism in particular would inevitably result in frequently losing the thread of my argument, for very little gain in real insight into that argument. It seems to me that there can – and perhaps ought to – be realistic hermeneutic fictionalism and realistic revolutionary fictionalism, and that the possibilities for each are illuminated by reflecting on the same sorts of issues, so far as the issues of Part I are concerned, anyway.

4. Fictionalism and Truth in Fiction

As I said above, in Sections 1 and 2, a very common way of thinking about fictionalism proceeds by means of a strong analogy between the fictionalist's target discourse and fiction: the idea is that in order to understand some target discourse we can/do (hermeneutic fictionalism) or should (revolutionary fictionalism) think of it more or less *as* a fiction. That initially suggests something like the following model: choose or discover a theory of fictions, and specifically of *truth* in fictions, and then apply that theory to the target discourse, analysing 'truth' – or rather appropriateness – in that discourse in the same way as 'truth' in fictions. This model quickly invites the thought that the claims of the target discourse are 'true according to' the theory or discourse, in essentially the same way as fictional claims (such as 'Sherlock Holmes was a detective') are true according to a story or plot (the Sherlock Holmes stories, or any particular one of them). This idea is available to fictionalists who reject the semantic replacement, operator approach to fictionalism I discussed and objected to in the previous section, too: whether or not the claims of the target discourse are to be *semantically* analysed as involving an *according to...* operator, they might be treated as true according to some fiction. (The point here is that one doesn't have to treat the *semantics* of fiction as, for example, Lewis, 1978 does in order to be entitled to the thought that there are some things which, *according to some fiction(s)*, are true.)

But things are not quite so straightforward, because for the analogy between the target discourse and fiction to be appropriate it must be the case (obviously!) that there is some relevant similarity between the target discourse and fiction (and our relation to it). And there is no immediate reason to suppose that the relevant similarities

exist, at least at the level of *philosophy* of fiction.¹⁶ At best, the analogy is a distraction since it threatens to import whatever theoretical commitments we have about the nature of fiction into discussion of the target discourse. At worst, it is positively threatening to the aims of the fictionalist who presses the analogy in the first place, since some commitments we might have about the nature of fiction simply wouldn't leave space for the advantages the fictionalist claims for their theory if imported into the analysis of the target discourse. I will briefly indicate two ways in which the analogy might break down.

If the apparent correctness of accepting claims such as that 'Sherlock Holmes was a detective' is cashed out (in the fictional case) in terms of some genuine application of a truth predicate, and if that truth predicate is cashed out in terms of something like reference to abstract fictional objects, then fictional claims will come out truth-apt (in line with fictionalism) but will plausibly *not* be *nondoxastic*. At least on one theory of the relation between readers and real fictional characters, if there is an abstract Sherlock Holmes — or an abstract set of 'Sherlock Holmes facts' — about which there are truths, my acceptance of those truths, in so far as it is determined by my reading of the stories, might well be quite straightforwardly *doxastic*: I just *believe* the claims about Sherlock Holmes — the abstract Sherlock Holmes — made in the stories. No *doxastic* theory of fictions could proffer an appropriate fictional analogue of the fictionalist's target discourse (since the fictionalist characteristically thinks that acceptance of target claims is *nondoxastic*), so on this theory of fictions, the analogy between fictionalism and fiction breaks down. (I do

¹⁶ In this section I am discharging my assumption that it is with our *ordinary, everyday* notion of fiction that the fictionalists' analogy is drawn, by exploring (rather briefly) what hope there is for drawing the analogy with some other notion of fiction.

not, of course, commend this theory of fiction to the reader; I merely note that someone *might* hold it.)

Similarly, it is not at all clear whether the best theory of fiction will support the everyday assumption that fictions are truth-apt at all. Indeed Frege notoriously thought that they are *not*, for fictional terms have sense but no reference.¹⁷ Whether or not we agree with Frege's reasoning, we might very well be persuaded, on reflection, that fictional claims are neither true nor false which, given bivalence at least, means that they are not truth-apt. We would then have a *non-representationalist* theory of fiction. And no *non-representationalist* theory of fiction could be a suitable analogue for a *representationalist* theory of some target discourse such as the fictionalist's. (Again, I do not commend this view, but merely note that it is available and perhaps viable, for all the fictionalist typically *proves* about the nature of fiction.)

Of course, depending on your theory of fiction, the analogy with fiction *might* be a good way of glossing fictionalism. But since some seemingly viable theories of fiction (see Friend (2007); and Sainsbury (2009) for a more thorough survey of options) turn out to be *incompatible* with fictionalism we should not assume that fictionalism is automatically a theory about its target claims *being* fictions, or even being *treated as* fictions. For this reason, 'fictionalism' might be an unhappy label, but we can retain it so long as we remember to be cautious of the common analogy.

5. Conclusion

¹⁷ See, e.g., Frege [1897]: pp229-30.

I have been attempting to say what it is about fiction that might be of use to a fictionalist. It does not seem that the typical falsity of fictional claims needs to be any part of what is essential to fiction at all. There seem to be quite unproblematic true fictional claims. If the fictionalist is concerned with our ordinary, pre-philosophical ideas about fiction, then they ought to be interpreted, I think, as being interested in the distinctively non-truth-directed acceptance norms commonly associated with fiction, for it is those norms which are plausibly central to the nature of fiction (on the everyday, pre-philosophical idea of fiction, at least).

Once we have seen that even on the 'everyday' analogy with fiction there is no reason to think that the fictionalist must treat their target claims as false, there is space for an intelligible sort of fictionalism that is fully *realist*. What I hope to have shown here is that the thought that there must be a problem with realistic fictionalism that comes from the very nature of fiction or make-believe is misguided. It is misguided firstly because even with the strong analogy in place there is no reason to think that the typical falsity of fictional claims is salient; and secondly because the strong analogy is not, on closer inspection, the most perspicuous way of saying what the fictionalist's commitments *are*, since their real motivations tend to be semantic and psychological theses which are strictly separable from the analogy and which ought to be separated so as to be clear about what is really at issue (which is presumably something about the target discourse, and not about what fiction happens to be like¹⁸).

If, in an attempt to illustrate the more theoretical semantic and psychological theses which underpin the original 'everyday' analogy, the fictionalist prefers to draw a strong analogy between their *theory* of some

¹⁸ Unless, of course, the target discourse *is* fiction, as it is for Brock, 2002, who proposes fictionalism about fictional characters.

discourse and the *theory* of fiction, it becomes a delicate question whether the correct theory of fiction is at all well suited to the fictionalist's purpose anyway, given the semantic and psychological theses to which they are committed. And that delicate question seems to be an unnecessary controversy for a fictionalist who is *really* interested not in the theory of fiction but in the correct theory for their target discourse. The fictionalist does better to simply put their theory in terms of the semantic and psychological theses themselves. And once attention is refocused on those semantic and psychological theses, it once again seems plausible that realism (as a metaphysical thesis) is a real option for one type of fictionalist. Jason Stanley refers to 'the spirit of hermeneutic fictionalism as a strategy for avoiding ontological commitment' (2001: p37). If that *is* the spirit of fictionalism, then it does not follow that any fictionalist who wants to avoid ontological commitment must be wanting that in virtue of rejecting realism. There are, perhaps, other, non-metaphysical reasons for wanting to avoid ontological commitment, reasons which we shall explore at greater length in Part II.

Fictionalism for Realists

1. Introduction

As we saw in the previous chapter, some fictionalists rely upon a distinction between *doxastic* and *nodoxastic* modes of acceptance: they argue that whilst we *believe* all sorts of claims about all sorts of things, we in fact accept claims of some target discourse (mathematics, possible worlds talk, the metaphysics of objects or universals, scientific theorising about unobservables, moral discourse, etc.) in some other way, or we ought to do so (depending upon whether they are ‘hermeneutic’ or ‘revolutionary’ fictionalists, respectively).

Some fictionalists of this sort already recognise that we *might* be fictionalists about some discourse even if some of an important class of the claims of that discourse are, or might very well be, *true*. As we shall see in Part II, there is more to be said for such realistic fictionalism than merely that it is *possible*. But suffice it for now to grant that there is some reason for some of us to be fictionalists of the realistic sort. Unlike most fictionalists, we do not argue for our fictionalism from an error theory, or an anti-realist or nominalist position about mathematical objects, possible worlds, or moral values, facts or properties, etc. We note that it is desirable to analyse the modes of *acceptance* of the target claims as the fictionalist does, and that it is desirable to analyse the *semantic content* of the target claims as the fictionalist does, but we also note that neither these accounts of acceptance nor of semantic content *commit* us to any non-realist *metaphysical* position. Indeed, we note that we are free to treat those claims which, if true, entail the existence of mathematical objects,

possible worlds, or moral values, facts or properties, etc as being true after all. So those of us who want to exercise our freedom within a fictionalist framework to regard some such claims as true are, in a respectable sense of the term, realists about numbers, sets, possible worlds, moral values or whatever.¹

In this chapter, I want to point out that a realistic fictionalist trivially avoids a recently pressed problem for fictionalism, a trilemma presented by Chris John Daly. But having explained this, a general objection to the realistic fictionalist's reliance on the thesis that one can *appropriately* both believe and 'nondoxastically accept' some claim will loom sharply into view. So I address that objection, which I call the Crowding Out Objection, by suggesting that we do not, in general, expect belief to crowd out nondoxastic acceptance in the way the objection assumes.

Of course a deeper worry about the possibility of *any* sort of fictionalism which depends upon a distinction between belief and nondoxastic acceptance turns on the plausibility of that very distinction in the first place. There are important issues about the *conceptual* plausibility of the doxastic/nodoxastic acceptance distinction, which I shall not address directly yet.² I address it properly in Chapter Three. The discussion I will present in this chapter does offer an *indirect* treatment of at least one form of the conceptual plausibility objection, though. It is

¹ See Yablo, 2002: sec.12, for some related discussion going in broadly the same direction. I shall not dwell on the issues raised by Yablo's discussion because he treats fictionalism as tied to treating the target discourse and our acceptance of its claims on a *pretence* model which, for reasons I discussed in Chapter One, I reject; and his discussion centres on the virtues of acceptance of the target claims being to do with the representational benefits of talking about a claim's 'real content' (as opposed to its 'literal content') in the language of the target discourse. As becomes clear below (see Section 5), I am not tied to the idea of fictionalism as having a special place for *representational* virtues, either.

² See, e.g., Horwich, [1991]; 2006. Daly, 2008 presents a number of such objections and defends fictionalism against them.

sometimes said that it is just by fiat that fictionalists introduce the doxastic/nodoxastic acceptance distinction, that it serves no explanatory or theoretical purpose other than to help the fictionalist out. But if the cases I am going to discuss are best characterized in terms of that distinction just because that is the most natural way of talking about them (as I think it is), then there seems to be at least one good *prima facie* reason to accept the distinction regardless of whether the fictionalists need it or not.

The Crowding Out Objection is what we might call a worry about the *practical* plausibility of the distinction when applied to the attitudes of the realistic fictionalist themselves: given that we can distinguish belief from nondoxastic acceptance, can someone who believes *P* (as the realist presumably must) also knowingly *nodoxastically* accept *P*? The crowding out objection says not, for full-blown belief crowds out other modes of acceptance, or at least crowds out the importance of other modes of acceptance. The objection can be met, I think, if we understand what it is to accept a claim in a variety of circumstances. It is not obviously true that belief crowds out other forms of acceptance, even if one knows that what is believed and nondoxastically accepted is the same (putative) fact.

2. Fictionalism

The sort of fictionalism which depends upon the distinction between doxastic and nondoxastic acceptance is best characterised, as I said in the previous chapter, as the novel conjunction of two theses, one about semantics and one about psychology.³ The semantic thesis is, I think,

³ My characterisation of this sort of fictionalism closely follows Kalderon, 2005a; 2005c.

quite familiar and variously goes under the names ‘descriptivism’ or ‘factualism’, and I am calling it ‘(semantic) representationalism’. It is the thesis that, literally construed, the fictionalist’s target claims purport to describe what there is, and how it is; and that they are therefore *truth apt*, their truth or falsity depending upon what exists and what it is like, and whether what exists and what it is like are as the target claims say they are. This thesis is the one notoriously denied (with respect to certain target discourses) by the emotivists and, in more subtle ways, by later expressivists.⁴ But for those who reject the expressivists’ semantic non-representationalism, the interesting thing about fictionalism is the other, psychological, or perhaps more properly ‘normative-psychological’, thesis.

That other thesis is what I am calling ‘nondoxasticism’, though care must be taken in formulating it.

Importantly, ‘nondoxasticism’ is *not* the thesis that the attitudes appropriate to some discourse are not *propositional* attitudes. It is a more restricted thesis about the mode of acceptance of the target claims appropriate to a particular practice or range of practices. Here is one attempt to formulate it, due to Daly (2008: p425), in which S is the target discourse:

⁴ See emotivists such as Ayer, [1936/46] and Stevenson, [1944], also R. M. Hare’s prescriptivism in, e.g., Hare, [1952], and expressivists such as Blackburn, 1984: ch.6, and Gibbard, 1990. Whether or not emotivists or expressivists need to deny truth-aptness, and what truth-aptness might be for them, depends upon what sort of conceptions of truth they are willing to countenance. Certainly Stevenson (1963b: sec.8; cf. 1944: ch7, sec.5) recognised a notion of truth (a use of ‘that’s true’) having to do with agreement in attitude, and Blackburn’s ‘quasi-realism’ is designed *inter alia* to earn the right to talk of truth (see Blackburn, 1993). What is clear, though, (as Stevenson acknowledged) is that truth-aptness for Stevenson or Blackburn will not be the same sort of reference-involving notion that it is for the representationalist (or ‘descriptivist’) for whom the truth of some claim depends straightforwardly upon whether that claim succeeds in saying how the things which exist really are. (See Rosen, 1994 for a nice discussion of issues to do with what phrases such as ‘saying how the things which exist really are’ might mean.)

- (1) The fictionalist does not believe any sentence of S, but
- (2) They believe that the sentences of S are useful.

But Daly's way of putting the fictionalist's point here is rather puzzling. Firstly, there is no reason to assume that the fictionalist's theory is directed at what the *fictionalist* does at all. Take Field's mathematical fictionalism: Field recommends that *mathematicians* (or physicists and applied mathematicians, at least) do not literally believe any of their ontologically committing mathematical claims, but that *they* ought to carry on regardless because those claims are useful. That is *not* to say that mathematicians must be fictionalists (it is sufficient for their mathematical practice that they accept the claims — they needn't reflect on their acceptance as the fictionalist *philosopher* of maths does). Fictionalism is a theory about how some participant in a target discourse goes about their business, or ought to go about it. It is an open question whether the *fictionalist* actually participates in the target discourse (not all philosophers of science are *scientists*, not all philosophers of maths are *mathematicians*...).

Better, then, to replace (1) and (2) with:

(1*) Those participating in S do not believe, or ought not to believe, any sentence of S;

But

(2*) The sentences of S are useful.⁵

⁵ This formulation also, in (1*), amends Daly's more limited formulation by adding the clause which makes it applicable to revolutionary as well as hermeneutic fictionalism.

This is still not the right formulation of the fictionalist's claim about *acceptance*, though. Suppose Bob believes that it would be useful to accept *P* even though he doesn't believe *P*. That does not yet mean that he has reason to accept *P*, since usefulness might not be related to an acceptance norm at all. ('Oh! If *only* I could accept *P*, for it would be *so* useful — but alas I cannot!') The fictionalist, distinctively, thinks that usefulness (or some other non-truth consideration) *is* related to an acceptance norm, but that the mode of acceptance to which considerations of usefulness are relevant is not belief (since belief is truth-, or evidence-, normed, not utility-normed). So in characterising fictionalism, we had better formulate nondoxasticism as a thesis about *appropriate modes of acceptance* of *P*, not just some belief about the usefulness of *P* (a belief about some property of *P* which might or might not have implications for whether to accept *P*).

Better, then, to replace (2) and (2*) with:

(2[†]) The appropriate mode of acceptance of sentences of *S* is some mode of acceptance other than belief; which appropriate mode of acceptance is not truth- (or evidence-) normed.

This formulation brings out two important aspects of the fictionalist's interest in the psychology of acceptance, which were at best left implicit in Daly's version: firstly, it emphasises that *belief* is not the mode of acceptance that the fictionalist is *normatively* interested in; and secondly, it introduces the idea of some norm — e.g., usefulness — that nonetheless governs acceptance of the target claims, in place of the usual norms of belief.⁶

⁶ (2[†]) makes explicit what is fictionalist about van Fraassen's (1980) 'constructive empiricism'.

However, this formulation of the nondoxasticism thesis is still not quite general enough to account for *all* the possible fictionalist positions. (1*) still has it that those participating in some target discourse do not believe, or ought not to believe, the claims of that discourse. Restrict what counts as the 'target discourse' so that only the ontologically committing claims of a more general discourse count. (That is, restrict the target discourse of mathematical fictionalism to only those claims quantifying, explicitly or implicitly, over sets, for example.) Still, (1*) is too strong to be maximally general. Some of those involved in even the restricted target discourse still might believe its claims on occasion, and might even be correct to. The target claims, as the *realistic* fictionalist is eager to point out, might, after all, be *true*. The fictionalist's most general psychological thesis, then, is better captured by:

Nondoxasticism

Appropriate acceptance of sentences of S is some mode of acceptance other than belief. (The acceptance norms for sentences of S are some norms other than truth- or evidence-norms.)

This formulation avoids denying that the fictionalist believes the target claims, even if the fictionalist is one of the relevant practitioners with the discourse: the most it claims is that belief is not *all there is* to acceptance of the target claims. (Think of it this way: *there is* appropriate acceptance other than belief, so belief norms cannot be the only norms governing acceptance for S.)

Of course, there are multiple modes of quite appropriate acceptance at work in all sorts of contexts involving all sorts of discourses which do not invite a fictionalist analysis. But the fictionalist need not claim that nondoxasticism is baldly *sufficient* for the target discourse

being apt for a fictionalist analysis (or even that nondoxasticism plus semantic representationalism is baldly sufficient for the same). The further condition which must be met in order for a discourse to be so apt is that the nondoxastic mode of acceptance must play some significant role in the analysis of that discourse. So, for example, a realistic fictionalist about mathematics would argue that our mathematical practice ought to proceed as the fictionalist recommends even though, as it happens, there are mathematical objects. That is, the *norms* governing mathematical acceptance are, according to the realistic fictionalist, *not* the norms of belief, even though true belief is possible and, sometimes, is actual had. This preserves the difference between a realist fictionalist and a straightforward realist, who acknowledges that there *are* nondoxastic modes of acceptance of mathematical claims, but thinks that in the end it is the norms associated with *belief* which ought to govern our mathematical practice. I will say that, for the fictionalist, non-truth-normed (or non-evidence-normed) acceptance is the 'central' mode of acceptance with respect to the target discourse. (Other modes of acceptance are possible and maybe even appropriate, but they can come and go as may be — what really *matters*, for the fictionalist, is nondoxastic acceptance.) And, as was implicit in my criticism of Daly's formulation of nondoxasticism, above, it is insufficient to diagnose appropriate and central modes of acceptance of target discourses *per se*; what is required is to diagnose appropriate and central modes of acceptance of claims of the target discourse *within a particular practice*. What is a central mode of acceptance for one practice with a target discourse may be entirely incidental to another practice with that same discourse.

(This point is important when we consider cases, such as the Moby Dick example, below. It has been objected to me that the fictionalist is saying nothing non-vacuous if all they are saying is that there are both

doxastic (belief) and nondoxastic modes of acceptance of some claims. This objection misses its mark, for – it bears repeating for emphasis – the fictionalist is concerned with the *appropriate* modes of acceptance (a normative notion), not just which modes of acceptance there are, or might be, and with which modes of acceptance are *central* to a particular *practice* involving the discourse.)

Fictionalists think that whilst the claims of their target discourse typically threaten to commit those who believe them to the existence of mathematical objects, or possible worlds, or scientific unobservables, or moral properties etc., believing those claims is only one of the possible appropriate ways of accepting them and, just as accepting that ‘Sherlock Holmes was a detective’ without believing it doesn’t commit us to thinking that there was really a Sherlock Holmes, accepting mathematical, or modal, or scientific, or moral etc. claims does not commit us to thinking that there really are mathematical objects, or possible worlds, or scientific unobservables, or moral properties etc. if our acceptance does not take the form of belief.

It is very clear why this sort of analysis is attractive to some non-realists: it offers an escape from the ontological commitments of various types of discourse, without denying that those discourses, literally construed, refer to or quantify over some ontological undesirables (abstracta, possible worlds etc.), and it achieves this by shifting the locus of ontological commitment from the literal claims themselves to the necessary commitments of those who *accept* the claims — best of both worlds for someone with a desire for a maximally spare ontology without a reductivist or elitivist semantics.⁷

⁷ This distinction between the ontological commitments of *sentences* or *propositions* and the ontological commitments of *people* was marked, of course, in the previous chapter.

But fictionalism also promises great things for those who think that the literally construed, ontologically committing, target claims are (generally) *true*, but who are moved by the various problems that realism has with its usual emphasis upon *belief*. Realists, for example, may be queasy about the *epistemological* issues raised by emphasis upon belief: they might think that our *beliefs* about their target discourse are *typically* (though not universally) lacking in warrant, and thus that in order for us to *typically* be allowed to go on accepting their target claims we had better be accepting them nondoxastically. That is, realists and non-realists alike can usefully avail themselves of the advantages of marking the distinction between the ontological commitments of some target claims themselves and the ontological commitments of those who accept them. But whilst the non-realist fictionalist makes use of this distinction by noting that the ontological commitments of their target claims are innocuous and *false* (innocuous because acceptance of them does not commit the careful practitioner to believing in the false ontology),⁸ the realistic fictionalist will say that there would be no error involved in practitioners with the target discourse being ontological committed, but that there remains reason to accept the target claims on some basis other than their truth, or the evidence for them (to accept them, that is, in accordance with norms other than those associated with belief).

Motivating such a realistic fictionalism is the work of Part II. My concern in the remainder of this chapter is with (i) the ability of any fictionalist (and, as it turns out, particularly the realistic fictionalist) to avoid Daly's Trilemma, and (ii) the stability of the realistic fictionalist's position (with respect to more or less any discourse) in the face of the

⁸ One big advantage here is that the fictionalist non-realist can avoid the usual non-realist's need to say that the price of avoiding excessive ontological commitment is having to do without using the target discourse, which discourse might be useful or have some other nonalethic virtue.

Crowding Out Objection, for if it fails to meet those challenges then whatever other virtues realistic fictionalism has will be all in vain.

3. Daly's Trilemma

Daly, 2008 has pointed out that fictionalists who are not careful to restrict the domain of their supposed nominalism face a trilemma: either they must retreat to Platonism (realism) about *something* such as meanings, or sentence types; or they must abandon their fictionalism and adopt a non-fictionalist form of nominalism at some point, in which case they might as well have been non-fictionalist nominalists all along; or their adoption of a 'mixed' theory threatens to make their theory look *ad hoc* and, to that extent, undesirable as a theory at all. The implication in Daly's presentation of his trilemma is that thoroughgoing fictionalism is in trouble, since thoroughgoing fictionalists *qua* thoroughgoing nominalists cannot satisfactorily carry out their project.

The trilemma is, avowedly, directed at the fictionalist's *motivation* (Daly, 2008: p433), and Daly assumes that the motive is the global rejection of abstracta. It is an open question whether or not particular fictionalists (even those who do indeed have nominalist inclinations with respect to their *target* discourse) subscribe to the program of doing away with *all* abstracta.⁹ But I will leave that issue aside. The point I want to

⁹ Most fictionalists are fictionalists about a *particular* discourse. Hartry Field, for example, says that his conclusion 'is not based on any general instrumentalist stratagem: rather, it is based on a very special feature of mathematics that other disciplines do not share' (Field, 1980: p vii; see also p2, on how Field is not committed to *any* universal claim *from* his fictionalism.) It is true that Field does not anywhere (that I know of) say anything to indicate that he wants to allow for any particular abstracta; but his arguments (see also Field 1989: esp. ch1) against the possibility of mathematical abstracta, or at least against the right of any metaphysics which is properly accountable to epistemology to assert the existence of mathematical abstracta (which are variants of Paul Benacerraf's argument against the possibility of knowledge of mathematical abstracta (in Benacerraf, [1973]), purged of the causal theory of knowledge upon which

make here is that a *realistic* fictionalist faces no such trilemma. This is because even with respect to his target discourse, he has no reason to be a nominalist *at all*. Realistic fictionalists will think that not only are their target claims, including perhaps the ones which literally construed purport to describe abstracta, representational and truth-apt; they are (sometimes) *true*. This means that what seemed to be a challenge to fictionalism *per se* is, if anything, a reason to prefer realistic to (global) nominalist fictionalism.

Note that I am not suggesting that Daly's trilemma counts as evidence for realistic fictionalism, for I think the trilemma is broken-backed already (since there is no reason to assume that even nominalist fictionalists are, *qua* fictionalists, *global* nominalists). I merely want to claim that *if* you are impressed by Daly's argument, you have no new argument against *realistic* versions of fictionalism and so, if you are drawn to fictionalism at all, you have some reason to entertain the realist version as, from the point of view of someone sympathetic with Daly, the only viable fictionalist option.

At the very least, and trivially, the realistic fictionalist has nothing to fear from the trilemma since they can insist that the first horn is unproblematic: they can avail themselves of the strategy of realism without fear of contradiction, and need not even wait until the issue of

Benacerraf's argument seems to rest) are not *presented as* general arguments against the possibility of any abstracta at all, whether or not Field happens to be drawn to global nominalism, and it is not clear that arguments of the same type would be as effective against, say, sentence types as abstracta as they are against mathematical knowledge as knowledge of abstracta. Similarly, van Fraassen (1980) argues for his constructive empiricism by noting that inconsistent models (both of which, in virtue of their inconsistency, cannot be strictly veridical) can do adequately well at representing the world in such a way as to generate empirically adequate predictions. It is not at all clear why such a motivating argument should be generalizable (so as to ground the global denial of abstracta) in the way Daly seems to assume. (It seems to me that perhaps Yablo, 2000 is the closest in spirit to the general attitude towards ontology and fictionalism that Daly's argument seems to require; which is not to say that it is close enough.)

meanings or sentence types arises to avail themselves of it — they are, after all, a realist about facts, objects or properties putatively described by the target discourse already.

But we ought to say that the realist fictionalist *accepts* that there are whichever facts, objects or properties (including, perhaps, abstracta) their realism commits them to, which acceptance must, on pain of regress, amount to acceptance in the *belief* sense: there is no reason for them to welcome a retreat to fictionalism about realism (according to which he nondoxastically accepts his realism), after all. Realistic fictionalists presumably think that their theory itself — including their commitment to realism — is literally believable (at least in the absence of an argument for why it is to be accepted at all if it is not to be believed, and there is no reason to just assume that any fictionalist about a discourse other than the debate about realism itself will be committed to any such argument). And this invites a deeper problem unique to the realistic fictionalist, the Crowding Out Objection.

4. The Crowding Out Objection

The crowding out objection concerns the coherence of the realistic fictionalist's own attitudes towards their target discourse and the modes of acceptance appropriate (and central) to it.¹⁰ They think that (some of) their target claims are true, but that our (central) appropriate acceptance of them is, or ought to be, some attitude other than belief. But, the objection goes, if the fictionalist themselves thinks that some target claim, *P*, is true then surely they must *believe* it, and think that they *ought to*

¹⁰ The Crowding Out Objection is due to Mike Martin and Moises Vaca Paniagua, who have pressed it independently of each other in various conversations with me about the coherence of realistic fictionalism. See also Yablo, 2002: pp82-3.

believe it. And if they believe P , and that they ought to believe P , then surely those attitudes of theirs towards P trump or crowd out any *nondoxastic* attitude's claim to be the central appropriate attitude to P , from their own point of view at least. After all, if you think that P is true, you had (plausibly) better *believe* P regardless of *whatever* other attitudes might be appropriate. Any attitude other than belief could, from your point of view, at best be *compatible* with the demand to believe P but could never *trump* that demand, so could not be the *central* mode of (appropriate) acceptance.

As is familiar by now, it is plausible that accepting that 'Sherlock Holmes was a detective' is correct. But if it *is* correct, it is presumably not in virtue of the usual *belief* norms, which surely include injunctions against accepting what you know to be false (which, knowing that there was no real Sherlock Holmes at least, we presumably do). According to the Crowding Out Objection, the belief norms are 'stronger' than the nondoxastic acceptance norms, so must be deferred to if one takes P to be true, leaving nondoxastic acceptance theoretically inert. Put another way, the objection is that once any fictionalist believes their target claim(s), there is just no *point* in insisting on the other appropriate modes of acceptance: the game is already up.

To illustrate the point of the objection, think about a mathematical realist who turns up to give a seminar. After several sessions, during which our realist has persuaded us that there *really are* sets (or functions, or numbers), the mathematical fictionalist at the back — who has been following attentively and has been thoroughly persuaded — raises his hand and says: 'I can see that there *really are* sets, so we should believe in them; but there's another reason for accepting — *nondoxastically accepting* — that there are, too: claims that commit us to the existence of sets are *really useful!*'. The intuition that the Crowding Out Objection trades upon

is the following. The mathematical realist would surely say to the fictionalist: ‘look, you are a realist now, you *believe* in sets; so why don’t you just say that the usefulness of claims that commit us to the existence of sets is just a nice bonus — the really important thing, the thing that really makes it important to accept that there are sets, is that there *are* sets, whether or not they’re useful, and the appropriate mode of accepting that there are things which you think there really *are* is belief!’.

5. No Crowding Out

But the Crowding Out Objection is too fast. There is, in general, no reason to think that nondoxastic acceptance is crowded out by belief at all. And, as noted above, fictionalism is not just the idea that there *is* some acceptance that is nondoxastic — it is the view that it is the noncognitive modes of acceptance that are somehow central.

Note, first, that there is no reason to think that, in general, we cannot consistently believe *and* nondoxastically accept the same claim, *P*, simultaneously. Consider the following case.

Suppose that there really was a captain Ahab who really hunted the white whale and that Melville knew of him and his adventures, elaborating them into the story of *Moby Dick*.¹¹ That there was a captain

¹¹ We know that Melville knew of a white whale, since he includes a passage from Harris Coll which speaks of such a thing in the ‘Extracts’ which precede the main narrative of the novel (‘One of our harpooneers [sic] told me that he caught once a whale in Spitzbergen that was white all over’, Coll apparently tells us in *A Voyage to Greenland, A. D. 1671* (Melville, [1851]: p. xviii)), and we know that Melville was also impressed by the account of a great whale attacking and sinking a whaling ship in 1820 given by Owen Chase, in a work also cited by Melville in the Extracts (Melville, [1851]: p. xxi). But I want us to suppose something more than that Melville, like more or less any novelist, was influenced by and incorporated various things he knew about his subject; I want us to suppose that Melville knew that there was the very white whale of which he wrote in *Moby Dick*, and that there was a captain Ahab of whom he wrote. These

Ahab, and that there was a white whale etc. are, we are supposing, facts. Now suppose that Bob is sitting his A Level exams: he has a Literature exam in the morning and a History exam in the afternoon. Answering a question on *Moby Dick* for the Literature exam, Bob writes down: 'There was a white whale, which was the object of Ahab's fascination'. He writes this down because he accepts that there was a white whale which was the object of Ahab's fascination on the basis of *Moby Dick's* plot. Of course, he does not, on most theories of fiction at least, literally *believe* it, because he accepts it only in so far as *Moby Dick* says it is so, and he does not consider *Moby Dick* a factual work: *he* does not know that Melville was aware of the white whale, or of Ahab. Bob does, however, know about the historical white whale and about Ahab, and plans to write about them in the afternoon's History exam where he hopes for the opportunity to write 'There was a white whale, which was the object of Ahab's fascination' because he literally believes that there was a white whale, which was the object of Ahab's fascination. His ignorance is of *Melville's* knowledge of the historical facts, not of the historical facts themselves — he thinks that it is interesting that *Moby Dick* comes so close to the historical facts, but it never occurs to him to think that what is reported in Melville's book *is* the historical facts. (Bob also, it so happens, thinks that it is remarkable how closely Orwell's *1984* resembles the subsequent historical facts, so he is quite familiar with 'accidental' coincidence of fact and fiction.)

Now *we*, supposing as we are that *Moby Dick* does report the historical facts (perhaps we are literary historians who have unearthed this truth about Melville's work in obscure documents which other readers have not seen), can see that a fact about which Bob will write in

suppositions go far beyond what (to my knowledge) was the case, for all of Melville's being influenced by the facts he knew of the world of whaling.

the afternoon (in his History exam) and a fact about which Bob will write in the morning (not believing it to be a fact, in the Literature exam) are *one and the same fact*. But we can also see, I think, that it is plausible that Bob has (on the day of his exams) two distinct attitudes to that fact: with respect to thinking about *Moby Dick* he does not believe, but does nondoxastically accept, a claim expressing the fact; with respect to thinking about maritime history, he does believe the same claim expressing the same fact (unknownst to him).

I am pretty sure this case is decisive with respect to the possibility of believing *and* nondoxastically accepting some claim (and with respect to the possibility of believing and nondoxastically accepting some claim *simultaneously*, for the case was constructed so as to focus on Bob's intention, in the morning, as he sits his Literature exam, to write about Ahab in the afternoon). Indeed, I think it is decisive with respect to the stronger claim that it can be *entirely appropriate* to believe *and* nondoxastically accept some claim (simultaneously), and even with respect to the still stronger claim that nondoxastic acceptance can be *central* to the analysis of some practice with the discourse even if it is belief which is central to the analysis of some other practice with the same discourse. But this will not quite do to overcome the Crowding Out Objection, since that objection trades upon the thought that the realistic fictionalist not only (i) believes *and* nondoxastically accepts some target claim(s), but that (ii) they *know* (or at least they *believe*) that the fact about which they both believe and nondoxastically accept some claims is the same fact. It is this knowledge, the objection goes, which makes it inconsistent for him to both believe and nondoxastically accept (or

recommend the nondoxastic acceptance of) the target claim(s) in question in any interesting sense.¹²

But even with that knowledge (or belief), there seem to be perfectly ordinary cases in which there is no crowding out, and where — even more significantly for the realistic fictionalist — nondoxastic acceptance remains the most important (central) mode of acceptance even with belief on the scene.

Suppose you are playing a game with a small child in the garden. A cat neither of you have seen before strolls into the garden and is incorporated into the game. ‘This cat,’ your small play-mate informs you, ‘is called Jeffrey’. Your play-mate doesn’t know what the cat is called, but in the game he has decided that the cat must be Jeffrey, and you, as a serious participant in the game, accept that he is called Jeffrey. You refer to him as Jeffrey, you think of him as Jeffrey while you play the game. As you go indoors to fetch another vital prop for the game, the phone rings. It is the lady who lives across the street, who says that she has lost her

¹² We can, if we wish, get a version of the Crowding Out Objection to apply to the realistic fictionalist’s recommendations for others’ nondoxastic acceptance of the target claims as much as to his own supposed nondoxastic acceptance of those claims. As we have seen, the supposition that the fictionalist themselves nondoxastically accepts the target claims (and that their nondoxastic acceptance is appropriate in a way apt to make it a candidate for being central to some practice for them) *might* be a supposition too far, since the fictionalist themselves might not be the sort of practitioner with the discourse whose acceptance of the target claims is subject to his their recommendations about nondoxastic acceptance (they might be a philosopher of maths, for example, but not a mathematician). But the force of the Crowding Out Objection for their recommendations of nondoxastic acceptance to others — even if they were no such practitioner themselves — would be to question whether the realistic fictionalist *could be* such a practitioner, or whether their realist commitments might stand in the way of their *possibly* adopting the attitudes they recommend (or at least their possibly adopting a *coherent* set of such attitudes), thus debaring them from engaging in the target practice as they recommend it be practiced. The idea would be that if *they* are debarred from coherently adopting the attitudes they recommend, they have failed to provide a recommendation which everyone — including themselves — could consistently satisfy, which plausibly ought to be seen as a failure of the recommendation. Happily, whether or not this form of the Crowding Out Objection is valid (and there is reason to doubt that it is), the realistic fictionalist is *not* debarred from adopting the attitudes they recommend anyway, as I go on to argue.

new cat which perfectly fits the description of 'Jeffrey' who is playing, or being played *with*, in the back garden. It turns out, surprisingly, that the cat's name is Jeffrey. You assure the lady from across the street that Jeffrey will be returned shortly, after some more play.

Returning to the game in the garden, you once again enter the world of the game. Now, of course, you actually believe that the cat is called Jeffrey. But when you refer to the cat as Jeffrey in the game with your young play-mate, when you think of the cat as Jeffrey in the game, you do so in just the same way you did before. It would be quite wrong to say that now, believing as you do that the cat really is called Jeffrey, you call him Jeffrey in the game because you believe it. No, you call him Jeffrey in the game because it was stipulated by your play-mate that the cat is called Jeffrey. The norm that is central to your practice here is the norm of accepting what the person whose 'game it is' says about such things as what the cat is going to be called. It is not the truth-interested belief norm, even though as it happens you can think of the cat as Jeffrey according to that truth-interested belief norm (since you have come to think that the cat is Jeffrey, from a reliable source). The belief norm is inert or redundant here, even though it is quite right to say that you do believe what you are nondoxastically accepting (you do believe that the cat is called Jeffrey, and *that the cat is called Jeffrey* is what you are nondoxastically accepting for the sake of the game, what you nondoxastically accepted — but did not believe — before taking the phone call indoors). Here, belief does not crowd out nondoxastic acceptance, nor does it crowd out the *importance* of nondoxastic acceptance. One way to see this is to ask whether you and the child are doing different things with respect to the name of the cat in the game. Clearly you are *not* doing different things: it is *not* that the child is going on according to the 'game-norms' whilst you, newly appraised of the

cat's real name, are doing something else when you refer to Jeffrey — you *both* refer to Jeffrey as Jeffrey *because* that is Jeffrey's 'game name'; *neither* of you refer to Jeffrey as Jeffrey because Jeffrey is his real name, even though as it happens you know that it is.

We can easily think of other cases in which full-blown belief does not crowd out nondoxastic acceptance. Many people think that they are subject to a norm of acceptance with respect to what their partner or loved-one earnestly tells them, according to which, roughly, you ought to accept what your partner tells you regardless of whether you would be justified in *believing* them or not — you ought, that is, to 'go along with' what they tell you (including making plans on the basis of it, replying in appropriate ways to questions from friends, etc...) whether or not you *believe* them, since your duty to abide by what they tell you extends beyond their epistemic authority (that is, you are obliged to act and think *as if* their word settled the issue of whether or not *P*, even though rationally it does not). Perhaps you would even be *irrational* to believe what they tell you on the basis of their testimony, because of some specifics of the case (of which you are aware) which make their testimony unreliable. All the same, your commitment to thinking and acting in ways that are sensitive to what they tell you (because you love them!) means you will accept (but nondoxastically accept) what they say.

Suppose you subscribe to such a norm. Now suppose that you have independent reason to believe that your partner was at the supermarket last night (you have come across their dated till receipt) when, as it happens, they *tell* you that they were at the supermarket last night (perhaps it just happens to come up in conversation as you are sorting through a draw and have come across the receipt). Plausibly, you end up with two distinct sorts of reason for accepting — in two distinct sorts of ways — that they were at the supermarket last night: you ought to

believe it, because you have evidence (which you have every reason to treat as reliable), but you also ought to accept it on the basis of their word, *even if* their word is not strictly reliable enough to ground *belief* but is, according to the acceptance norm just sketched, sufficient to ground some other form of acceptance. If we think that, when pressed about why you accept — note: why you *accept*, not specifically believe, or nondoxastically accept — the claim, it would be right for you to say ‘because they told me’, then nondoxastic acceptance is not only not inert but is seemingly the *central* mode of acceptance (it seems to be doing the explanatory work, seems to be the reason you accept the claim). And it seems it would be right to say that: wouldn’t you be suspicious of someone who seemed like they valued hard evidence over the assurances of their loved ones? The point here is, of course, precisely not about belief formation. Of course the testimony of someone unreliable is not a good way of coming to form beliefs. But the point is that the truth-directed belief norms are not all there is to whether or not there is reason to *accept* some claim. In cases such as these, it seems that belief does not crowd out nondoxastic acceptance. The nondoxastic acceptance norms, that is, have a real role to play in explaining the correctness of accepting some claim in these cases.

Note that, in this case, the norm associated with nondoxastic acceptance is not a utility norm (notwithstanding that life might go more smoothly if you accept what your partner tells you!), but rather a norm to do with what is appropriate with respect to dealing with a loved-one. Similarly, the acceptance norm in the previous case was a norm of make-believe. This does not at all undermine the point, though: fictionalism need not insist that it is only utility norms that ground nondoxastic acceptance. As we saw in the previous chapter, the spirit of fictionalism is not tied in any essential way to analogies with fiction or with make-

believe, imagination or any other domain, even though such analogies are very often employed. And likewise, fictionalists are not bound to think that what makes their target discourse apt for nondoxastic acceptance of its claims is something to do with *utility*, albeit that many fictionalists (particularly mathematical fictionalists (e.g. Field, 1980; 1989; Papineau, 1988; Yablo, 2005), scientific fictionalists (e.g. van Fraassen, 1980), and *some* moral fictionalists (Joyce, 2001; 2005)) do, in fact, propose utility as the grounds upon which to build their nondoxasticism. It is not clear whether a fictionalist about fictions (or, more accurately, about fictional characters), such as Brock, 2002, has much hope of persuading us that the reason to nondoxastically accept some fictional claim is the *usefulness* of that claim. And the fictionalist theory of ordinary moral deliberation I shall develop in Part II does not advert to the utility of our moral commitments (indeed I argue, in Chapter Six, that there is at least one reason to reject a model of ordinary moral commitment which accords approval only to the instrumentally useful commitments.)

The fact that doxastic and a nondoxastic attitudes towards the same propositions can sit side by side is not unfamiliar in the psychology literature on imagination, because of work done by Alan Leslie on children's games involving imagination. (See Leslie, 1994) It turns out – perhaps unsurprisingly – that children participating in a game which involves imagining that one cup is full whilst another is empty, when in fact both cups are empty, will reliably present the 'empty' cup as part of the game, showing that their imagining it as empty is getting them to differentiate it from the 'full' cup, and that they believe, of each cup, that it is empty, thus manifesting both the belief that the 'empty' cup is empty and the belief that it is empty.

Philosophical applications of these psychology results to issues in the philosophy of fiction have sometimes focussed on the way in which

they seem to support the idea that some particular theory about cognitive processing (such as the ‘single code hypothesis’) can help with familiar issues to do with our empathy towards fictional characters, for example. (See, for example, Nichols, 2004.) But other philosophers working on imagination and fiction have employed those results in ways which are more interesting for our purposes here. As Stock, 2011 points out, the Leslie results seem to imply that there is no incoherence in *prescribing* imagining what you know to be true. Stock is interested in the idea that the right way to think about fiction is as a prescription to imagine, and she is concerned to argue that we need to get clear about what imagination *is* before we can explain the Leslie results in such a way as to shed light upon the questions many have discussed about true fictional claims. Since our interest is not with fiction, though, but rather with *fictionalism*, and since we have seen that fictionalism need not be tied to any thesis about the fictionalist’s target discourse being treated as a fiction, or being treated in any particular way in respect of make-believe, pretence or imagination, the point of relevance for us in what Stock points out is that since it familiar from such everyday phenomena as children’s games of imagination that doxastic and nondoxastic attitudes often coincide, it would be extremely revisionary to say that there is some problem with *recommending* or *approving of* such coincidence – that would require us to say that there is something deviant in such everyday practices which there plainly is not.

I am not claiming — and I do not need to claim — that at in general nondoxastic acceptance crowds out belief any more than belief crowds out nondoxastic acceptance. It seems to me that in general there need be no presumption of crowding out at all when there are both doxastic and nondoxastic norms playing a role in someone’s acceptance of some claim for different purposes. Perhaps the way I have presented

the cases suggests that I think nondoxastic acceptance is crowding out belief. But in fact all I have tried to show is that there can be cases of belief and nondoxastic acceptance side by side, and further that there can be cases where nondoxastic acceptance is playing more of a role in our practice (a more central role) than belief is. Belief can play *some* role, though — for example, in the first case, in responding to the child if he asks later whether they had given the cat the same ‘game name’ as his actual name. So it is misleading to say, at the level of the target *discourse*, that the centrality of the nondoxastic modes of acceptance is crowding out belief, since all that is required for belief to play a central role with respect to the discourse is to move to a different practice with the same discourse (the practice of subjecting the game claims to a ‘truth test’ as opposed to just using them in playing the game, for example). At most, the crowding out of centrality might happen *within particular practices* with a target discourse; but, as I hope to have shown, there is no reason to assume that in general it is *nodoxastic* acceptance whose centrality to a practice with a discourse is crowded out.

It might be that nondoxastic acceptance is *central* to appropriate mathematical, or scientific, or moral practice, even if belief in particular claims is *philosophically* appropriate (because the claims happen to be true, though their literal truth is of no interest to mathematicians, scientists, moralists etc.).

If there are more specific crowding out worries about particular target discourses and practices, let those who press them attempt to do so in a way which doesn’t beg the question against the fictionalist, who in developing their position on some particular topic will presumably have a story about the virtues of nondoxastic acceptance, and the importance of the norms governing it. It certainly does not seem that *in general* there

is a crowding out phenomenon of the sort that would undermine the realistic fictionalist's position.

If more specific Crowding Out Objections are to be developed, they must be careful to steer clear of the mistake I noted in Section 2 when repairing Daly's characterisation of fictionalism: there is an important distinction between discourse and practice, such that (for example) philosophers of maths and mathematicians can be involved in very different practices (namely theorising about the ontology of mathematics and just *doing* maths) though they are concerned with the same discourse (mathematical claims). This is just as true of an individual pursuing different interests in a discourse, such as when a philosopher of maths who is also a mathematician moves between doing maths and doing *philosophy* of maths. Any serious Crowding Out Objection to realistic fictionalism for mathematics (for example) must be careful to allow that norms of acceptance are likely to be different for different practices (such as maths and *philosophy* of maths), and crowding out 'between' practices looks unlikely. A realistic fictionalist who recommends (or diagnoses) nondoxastic acceptance of mathematical claims for the pursuit of mathematics (or of physics, or engineering) need not be embarrassed by fact that with their *philosopher's* hat on they also believe those mathematical claims, literally construed, so long as their *mathematical* practice is governed by the norms they recommend. That is not to say that the Crowding Out Objection must be met by saying that the fictionalist must forget their belief in the mathematical claims when they don their mathematician's hat; it is just to say that their belief (which they might very well keep very much in mind) is *incidental* to what they are doing when doing mathematics (or physics, or engineering).

6. Conclusion

Fictionalists can be realists too. And besides whatever other advantages there might be to this strange-sounding conjunction of views, realistic fictionalists are particularly well placed to avoid Daly's Trilemma (indeed, they avoid it trivially). But if a fictionalist is to maintain realism about some discourse, they must surely *believe* at least some of their target claims, literally construed, which they recommend (if they are revolutionary fictionalists) we *nondoxastically accept*. Isn't this an uncomfortable position? No, because belief does not crowd out nondoxastic acceptance, and nor does believing the target claims mean that the *norms* of belief (though satisfied, if the belief is appropriate) are the most important norms for determining which claims to accept — it might still be the norms governing nondoxastic acceptance that play that role (*for a particular practice in question*).

The Varieties of Acceptance

1. Introduction

As is familiar by now, fictionalism is the conjunction of semantic representationalism and nondoxasticism. Representationalism is a semantic thesis about the content of the target claims: it says that their semantic content consists in — and is exhausted by — the stating or describing of putative facts. ‘Facts’ here are not metaphysically contentious: a ‘fact’ in this context is just a way the world is. It stands opposed to non-representationalist semantics for the target discourse, which hold that the content of those claims is at least in part determined by some non-(putative)-fact-stating element (such as what Fregeans call ‘tone’, or ‘colouring’ – see, e.g., Copp, 2001). An expressivist semantics for moral discourse, for example, aims to circumvent the apparent implausibility of there being real moral properties out there in the world by re-construing moral claims as claims with a distinctive *use* or *function*, which use or function contributes to determining semantic content. Representationalist semantics, on the other hand, does not deny that the target claims *have* a use or function — they can be used, at the very least, to describe the world — but denies that the use or function ‘gets into the semantics’.

The other, psychological or epistemological, thesis characteristic of fictionalism is concerned with the appropriate mode of acceptance of the claims of some target discourse. Specifically, ‘nodoxasticism’ is the thesis that the appropriate mode of acceptance of the target claims is some attitude other than *belief*. I am calling those attitudes which are not

belief 'nondoxastic' because 'mere acceptance' implies that the attitudes in question somehow fall short of belief, whereas I want to allow that they might be perfectly 'full blown' attitudes in their own right, and are not *lacking* with respect to any comparison with beliefs. I also prefer not to speak of 'noncognitive' attitudes, or of 'noncognitivism', because I want to leave it open whether the attitudes in which the fictionalist is interested are attitudes of a cognitive sort. How to draw a distinction between cognitive and noncognitive is not at all clear, but it seems unlikely that any plausible way of drawing it will put all and only beliefs on the cognitive side and the nondoxastic attitudes I am interested in on the other.¹

The fictionalist's distinction between belief and nondoxastic acceptance is contentious, and has come in for criticism. Critics of fictionalism have often complained that the fictionalists' positing of distinct states of belief and acceptance is at best *ad hoc* and at worst mistaken. (See Daly, 2008, for a survey of some such criticisms, and a defence on behalf of the fictionalist against them.) The aim of this chapter is to begin to develop a theory of acceptance which is safe for

¹ Here, in these choices of terminology, is one, relatively unimportant, way in which I differ from Kalderon, 2005a, whom I have tended to follow in my formulation of fictionalism as the conjunction of these semantic and psychological theses. Kalderon favours the nice slogan 'noncognitivism without nonfactualism' to sum fictionalism up, because he calls the sorts of expressivist semantics to which the fictionalist is opposed 'nonfactualism' and the psychological thesis I am calling nondoxasticism 'noncognitivism'. I favour 'representationalism' over 'factualism' simply for the reason that it is less counterintuitive (and possibly confusing) to call fictionalists 'representationalists' than to call them 'factualists', something I think Kalderon agrees with and explains why he speaks of '*nonfactualism*', but never of 'factualism'. It is quite natural for Kalderon to label the other thesis 'noncognitivism', because on the broadly Scanlonian moral psychology developed in Kalderon, 2005a: ch1 the nondoxastic attitudes attributed are, plausibly, not just nondoxastic but noncognitive. Classifying fictionalism as a species of psychological (not semantic) noncognitivism is very common. I seek a more generally applicable terminology, though, which remains applicable to theories according to which the attitudes concerned are nondoxastic but still characteristically cognitive. See Papineau, 1993: ch6 for a discussion of fictionalism which also employs the terminology of 'nondoxastic' attitudes.

fictionalism. This will involve (i) characterising acceptance in the most general way possible and (ii) characterising belief in such a way as to distinguish it from nondoxastic acceptance. This done, it will be possible, I think, to overcome the charge of misguidedness against the fictionalists' doxastic/nodoxastic acceptance distinction, and I will argue that at least one particular misguidedness objection (due to Horwich) is implausible. But in order to overcome the *ad hocery* objection it will be necessary to have (iii) developed a theory of acceptance which is independently plausible, regardless of the fictionalist's purposes.

The other thing I shall attempt to do here is to begin to make clear the distinctive way in which the norms which the fictionalist recommends in place of the truth-norms² associated with belief are supposed to work, and why those norms ought to be thought of as attaching particularly to appropriate *modes of acceptance*. But first we need to know what we mean by 'acceptance' and 'belief'.

2. Acceptance and Belief

What sort of state is a person in if they 'accept' *P*? And is there any way in which they might 'accept' *P* but not 'believe' *P*? At least since Jonathan Cohen's work on belief and acceptance (Cohen, 1989; 1992), there has been an influential strand of thinking about these questions — on both sides of the debates — which takes the putative distinction between believing and accepting to be primarily a matter of the *active*, voluntary

² For my purposes here, it does not matter what the precise form of the truth norm of belief is, nor whether, in fact, the appropriate norm is a truth norm or a truth-associated norm, such as an evidence norm. (Evidence norms are 'truth-associated' because evidence that *p* is always evidence that *P* is true; and the value of evidence here is linked in an essential way to that connection with truth.)

nature of acceptance versus the *passive* nature of belief.³ (See, e.g., the papers collected in Engel, 2000.) The spirit of Cohen's suggestion is summed up thus:

In my sense to accept that p is to have or adopt a policy of deeming, positing, or postulating that p — that is, of going along with that proposition (either for the long-term or for immediate purposes only) as a premises in some or all contexts for one's own and others' proofs, argumentations, inferences, deliberations, etc. Whether or not one assents and whether or not one feels it to be true that p . (Cohen, 1989: p368)

'Assenting' or 'feeling' it to be true that p is what it is, for Cohen, to *believe* that p . For Cohen, the answer to our second question is 'yes': a person might accept that p without believing that p , since one might employ a premise in an argument, for example, which one does not believe (or 'feel to be true') since 'the reasons for accepting that p need not always be epistemic ones: they might be ethical or prudential' (1989: p369). Similarly, a person can believe that p without accepting it:

[H]e could be convinced that p while nevertheless rejecting the use of that proposition as a premiss for any proofs, deliberations, etc. For example, he might be given a highly confidential piece of information after swearing that he will never rely on it as a premiss in any of his reasonings, even in the privacy of his own mind. (1989: p369)

Cohen's way of characterising belief and acceptance by appealing to aspects of the voluntariness of acceptance contrasted with the (relative) passivity of belief (the fact that believing that p doesn't typically require *doing* anything, even that belief is a type of *feeling*), and his view that neither entails the other, has been neither uncontroversial nor the only attempt at drawing the belief/acceptance distinction along similar lines.

³ Whether beliefs are to be treated as involuntary as well as passive is another question. Certainly it is often thought that beliefs *can* be voluntary, that we can decide to believe, and that doxastic voluntarism accords nicely with our tendency to hold others and ourselves responsible for our beliefs (see, e.g., Ginet, 2001). This does not mean, of course, that beliefs *cannot* be *involuntary*. Cf. Bennett, 1990.

(See 'Introduction' to Engel 2000 for a brief overview of alternatives and criticisms.) But whether or not any of the technical objections to Cohen's theory work, there seems to be a *taxonomical* problem with it. Specifically, there seems to be no principled way, on Cohen's taxonomy as stated so far, of distinguishing *acceptance* from *assumption*. According to Cohen's way of characterising what it is to accept *P* (the proposition *that p*), it is sufficient for accepting *P* that I have a policy of employing *P* as a premise in an argument. But suppose I only have such a policy of employing *P* as a premise because I think the best argument for not-*P* takes the form of a *reductio* from *P* to contradiction, and I think that not-*P* is the right result, so I have reason to employ such an argument whenever faced with a defender of *P*.⁴ Surely, on any natural understanding of 'acceptance' other than Cohen's technical notion, it is not right to say that I *accept P* in such a case — my assumption that *P* (for the purpose of *reductio*) is explained at least in part by my *rejection* of *P*, my desire to show a defender of *P* that he is mistaken.

There is actually a deeper problem here than just the implausibility of saying that I accept *P* just in virtue of employing it for the purpose of *reductio*. It will not do to say just that I don't *believe* (but do accept) *P* here, at least not on Cohen's model: there is a problem with *reductio* even if we allow that I don't believe but do accept *P* for such purposes. That's because Cohen thinks that accepting *P* commits us to accepting whatever follows from *P* (1989: pp370-1). The problem with the *reductio* case is that if my assumption that *P* counts as acceptance, then I must (if Cohen is right) be committed to not-not-*P* (assuming classical logic, at least); but if I am committed to accepting not-not-*P*, then in accepting not-*P* at the

⁴ For the purposes of this point I think we can read 'think that...' as meaning either 'believe that...' or 'accept that...', even on Cohen's understanding of belief and acceptance.

conclusion of my argument *I* am being inconsistent (since not-not-*P* contradicts not-*P*). There is a sense in which the very point of *reductio* is inconsistency. But whatever sense that is must be to do with the inconsistency of *premises*, not with my inconsistency as a reasoner! Of course we are often inconsistent in the sense of accepting contradictory propositions. What is implausible is that we might be committed to such contradiction just in virtue of going in for *reductio* style arguments (for conclusions about which we are not at all confused). There is a difference between saying that we are rationally committed to accepting whatever follows from *P* and saying that we are rationally committed to rejecting what is inconsistent with *P*. But even if Cohen's official theory says nothing about being committed to rejecting what is inconsistent with *P*, it remains the case the *I am* being inconsistent if I am committed, in a truth interested way, to both *P* (or not-not-*P*) and not-*P*. And it is just very implausible that my status as a competent (i.e. consistent) reasoner should be called into question by my employment of an argument form as seemingly unproblematic as *reductio ad absurdum*.

This mistake of reading the inconsistency of a *reductio*'s premises into the rational competence of anyone who employs such an argument is, I think, a reason to reject Cohen's characterisation of acceptance. But there is another, similar, problem. If I were to count as accepting the conclusion, not-*P*, I would, in virtue of that acceptance plus my supposed acceptance of the premise, be committed to accepting *P* and not-*P*, and since anything follows from a contradiction I would be rationally committed to accepting *any* (and every) proposition. That just cannot be right — I surely cannot be rationally committed to accepting just any proposition *just* in virtue of carrying out a *reductio* argument. So if acceptance commits me to consistency, as Cohen thinks, I either cannot properly accept the conclusion or cannot properly accept the premise in

question. I think it is clear which way we should go: we ought to say that I properly accept the conclusion and *don't* accept (but merely *assume*) the premise. To go the other way would completely falsify the only reasonable account of my reasons for giving the *reductio* argument at all.

A friend of Cohen might want to avoid this objection by revising Cohen's theory, jettisoning the controversial claim that accepting *P* commits us, rationally, to accepting whatever follows from *P*. But if we were to weaken the role of rational commitment in Cohen's account of acceptance, it would be even less clear than it already is what reason we would have for treating acceptance, *a la* Cohen, as the sort of thing that is *essentially* to do with reasoning, deduction etc. In any case, Cohen recognises the general consequence of saying that accepting contradictory propositions involves 'unwitting' commitment to every proposition, and bites the bullet whilst suggesting that 'there is a sense of "accept" in which acceptance is not thus deductively closed' (1989: p371). So let Cohen's technical notion of acceptance stand, understood now (as he apparently intends) as saying that:

A accepts that *p* iff (i) *A* adopts a policy of employing *P* in his reasoning [the original condition] and (ii) *A* is committed (wittingly or unwittingly) to accepting any proposition *Q* that follows from *P*.

If this is Cohen's technical notion of acceptance, it is clearly not the notion of acceptance we want for fictionalism, for no fictionalist has any reason to think that mere acceptance is closed under entailment. So there is reason to seek a different characterisation of belief and acceptance for the fictionalist's purposes. That is not to say that Cohen's (properly amended) theory of acceptance is wrong, for as we have seen Cohen is happy, in the end, to allow other senses of acceptance aside from his

technical one – perhaps *that* is the notion the fictionalist wants, and which is independently plausible. In any case, if the characterisation of acceptance and belief the fictionalist needs is to be independently plausible enough to avoid the charge of being merely *ad hoc* it will need to be more than just another technical notion of acceptance (and belief) to put alongside Cohen's.

This requirement can be satisfied, however. It would be satisfied if the fictionalist's theory of acceptance were maximally general — that is, if it accounted for all the cases in which we want to call an attitude 'acceptance' (unlike Cohen's, for he explicitly allows for other senses of 'accepts'), whilst still explaining why assumption (for example) is not acceptance. Such a maximally general theory would not be *ad hoc*, for it would serve as a theory of the state (and its varieties) a person is in when they are in the sort of state we talk about all the time. So it would be of general philosophical interest, and not just a technical device for the fictionalist's purposes.

3. A Theory of Acceptance

Belief is a mode of acceptance.⁵ There are other modes of acceptance, too. I am lumping all those others together and calling them nondoxastic acceptance. But what is 'acceptance'?

We might hope to appeal to some dispositional account: *accepting* *P* amounts to being disposed to $\Phi_1, \Phi_2, \Phi_3, \dots$. Perhaps the idea is that *accepting P* is being disposed to *assert that p*, or *assert P in particular*

⁵ Why think of belief as a mode of acceptance, and not just as something different from any sort of 'acceptance'? Some fictionalists, certainly, seem to adopt the latter approach. But I am inclined to suspect that the latter approach rather fetishises the *terminology* of belief and acceptance talk. It seems very obvious to me that we often say and think that a person accepts something they believe just in virtue of their believing it.

contexts. But dispositional accounts of anything inherit the difficulties of knowing what to make of dispositional claims themselves, and in any case it is not at all clear what sort of non-circular dispositional account could be offered for acceptance: plausibly the analysis of *asserting P*, for example, will appeal to *accepting P* so that appealing to *asserting P* in the analysis of *accepting P* will be problematic, or at least controversial. I think we can say something substantial about acceptance that is far less controversial than any dispositional account would need to be.

A sophisticated strategy can be found in Gibbard (1990). Gibbard, addressing the question ‘What is it to accept a norm?’,⁶ appeals to ‘tendencies’ to avow the norm. (Is ‘tendencies’ just another word for dispositions? Plausibly not: I tend to shop in Sainsbury’s, but am I thereby *disposed* to shop in Sainsbury’s? A tendency might just be a regularity, it need not involve the idea of a dispositional property.) One comes to tend to avow the norm in discussions about what to do, or about right and wrong. But that tendency to avow the norm is not yet to *accept* it, for we avow all sorts of norms which we do not accept for all sorts of reasons (deception, humour etc.). If, however, I conform to the demand for *consistency* in my avowals of norms — Gibbard thinks of this demand for consistency as a discursive demand, but other models could make use of other sources of such a demand — then I count as accepting those norms. Accepting a norm is ‘taking a position’, and part of that is adopting a *stable* position. That is what marks the difference between ‘accepting’ and just ‘internalizing’ a norm: tending to avow that norm is constitutive of both attitudes, but being responsive to the demand for

⁶ Note: accepting norms (the accepting *of* norms), not acceptance norms (the norms *governing* acceptance) which I will be concentrating on later. Gibbard does, though, think of what I am calling ‘demands’ as norms, so in his model normativity comes in at two levels: the acceptance *of* norms, and the norms *governing* normative discussion and commitment. I do not pretend to do justice to the breadth or depth of Gibbard’s view here, of course.

consistency (adopting a stable position) is characteristically constitutive of acceptance. (Gibbard, 1990: pp71-6)

Whether the same account — or a recognisable modification of it — will do for acceptance *in general*, though, is another matter. It is not clear, for example, what the demand analogous with the demand for consistency might be for accepting that Sherlock Holmes was a detective. Presumably we would not think of it as a transgression of the demand for consistency to accept (nondoxastically accept) that Sherlock Holmes was a detective *and* to accept (believe) that Sherlock Holmes was *not* a detective (because he doesn't – and didn't – exist). In fact that seems to be precisely the correct conjunction of attitudes to adopt, for it reflects the way in which it is correct to say 'Sherlock Holmes was a detective' and that 'Sherlock Holmes didn't really exist'. Whether or not there *really is* a demand for consistency here, *being responsive to* the demand for consistency is not plausibly necessary for acceptance in such cases⁷ — if there is a genuine demand for consistency here then there is reason to criticize someone who adopts both these attitudes to Sherlock Holmes, but feeling the force of that demand is not plausibly part of *what it is in general* to accept something about Sherlock Holmes, even if it *is* constitutive of accepting *norms* that one is sensitive to such demands.

What we are after is the most general characterization of acceptance. So perhaps a less reductive approach to thinking about acceptance in general would be useful. Specifically, it might be profitable to stake out some of the important or interesting *features* of acceptance with which we are familiar. The suggestion is that we pursue a project

⁷ I take it that being responsive to the demand for consistency in Gibbard's sense is not just a matter of whether one's avowals *are* consistent, but that 'being responsive' to a demand is just as much to do with *taking responsibility for* meeting that demand, trying to meet it, taking a dim view of not having met it (in relevant contexts) and the like.

similar to what Wright has recommended for thinking about truth (see, e.g., Wright, 1999: esp. sec.VI).

What we are after is a set of what Wright calls ‘platitudes’ about (in our case) acceptance, satisfaction of which is distinctive of acceptance. Like Wright’s truth platitudes, the features of acceptance to which we will be adverting are not uniquely realized. But modes of acceptance will be individuated by their realization of further features, or perhaps their different ways of realizing the common features. So we would, if we pursued the strategy relentlessly, end up with a taxonomy of modes of acceptance sufficiently fine grained to account for all the various modes of acceptance that we commonly recognise, including belief, the unity of which would be accounted for by their common realization of the most general features of acceptance and their individuality by their realization of various additional features or their different ways of realizing the common features.

I will not pursue that strategy *relentlessly* here, though. I will settle for suggesting what I think is the most important general feature of acceptance, and then a further feature of belief. If that further feature is indeed distinctive of belief, if it is a feature of belief but is not a feature of other modes of acceptance, then the distinction the fictionalist needs between belief and nondoxastic acceptance will be vindicated.

Just how general does our most general characterisation of the features of acceptance need to be? It does not, I think, need to be so general as to account for all the talk of ‘acceptance’ that goes on. If I say, for example, that ‘I have finally accepted that Tibbles is dead’ I might be saying something about whether I think that my cat is dead or not (I might be saying that, Tibbles having been missing for so long, I have now drawn the conclusion that he is dead). But I might be saying something quite different: I might be saying that I have *come to terms* with Tibbles

being dead. If that is what I am saying when I claim to have ‘accepted’ that Tibbles is dead, then I probably already believed that he was dead before ‘accepting’ it in this sense. Such uses of ‘acceptance’ are not of even the most general sort that we are interested in, for *whatever* we say about accepting *P* in the sense that the fictionalist means it, it will always be a further question whether we ‘accept’ *P* in this other sense. Indeed, it makes no sense to ask whether someone ‘accepts’ *P* in this other sense if one doesn’t already think that someone meets some general conditions for accepting *P* in the fictionalist’s sense — it would be silly to say ‘I accept that Tibbles is dead, but I don’t think he is’ if ‘acceptance’ is meant in this *coming to terms* way. (We can certainly make sense of ‘I accept that Tibbles *might* be dead’. But of course that is not an attitude towards the death of Tibbles; it is an attitude towards the possibility or likelihood of the death of Tibbles. And again, if I am to come to terms with the likelihood of Tibbles being dead, I must already accept it (the likelihood). We might also be able to make sense of ‘conditional acceptance’: ‘If Tibbles is dead then so be it’. But isn’t that elliptical for something like ‘since Tibbles might be dead, so be it’, which is just having come to terms with the likelihood or possibility of Tibbles’ demise?) The fictionalist’s general notion of acceptance is prior to this sense of ‘acceptance’. But this does not threaten the promised generality of the theory: the ‘coming to terms with’ sense of acceptance is sufficiently different from the sense in which acceptance is even putatively identical with belief that it deserves to be treated as a mere homonym, albeit one whose sense is not *entirely* dissociated from the one in which we are interested, in ordinary talk and not just by philosophical *fiat*.⁸

⁸ The same cannot be said for Cohen’s other senses of ‘accept’, which seem to differ in sense (in some cases at least) only in so far as they satisfy or fail to satisfy a quite technical formal condition (deductive closure).

One of the most important general features of accepting *P* (in the relevant sense) is that doing so puts one in a position to see thoughts or utterances of *P* as *correct* or *appropriate*. So, if I accept that *p*, asserting *P* will — in the relevant circumstances — seem like the correct or appropriate thing to do. Or, if I read or hear *P* it will strike me that what I have read or heard is correct or appropriate in some way.

When we *say that P*, for example, we are doing something. We might be *asserting P*, in which case what we might be doing is something to do with manifesting what we think is true. But we might be doing all manner of other things, many of which we might call ‘quasi-assertion’ (see Kalderon, 2005a: pp119-29). We might be *pretending*. Or we might be ‘asserting’ or ‘quasi-asserting’ something fictionally, as we do when we write ‘Sherlock Holmes was a detective’ in a literature exam. In the exam, as in other contexts, I need not preface that sentence with ‘according to the Conan Doyle novels’. Indeed, I might doubt that Conan Doyle wrote the novels, so it need not even be implicitly asserted that ‘*according to the Conan Doyle novels* Sherlock Holmes was a detective’. I may not even have any determinate idea of what the source of the stories is, whether they are novels or folk stories, written down or oral, whether there is a particular version or many variations. So I might not be (even implicitly) prefacing my answer with *any* ‘according to...’ operator. But nonetheless I am not asserting that Sherlock Holmes was a detective in the usual way or for the usual reasons. Neither, though, am I necessarily pretending that Sherlock Holmes was a detective: I may have no interest at all in ‘entering into’ the fiction in the sort of way that pretence seems to involve. Standards of correctness vary depending upon what is being done: saying that *p* might count as correct if my saying that *p* is an ordinary

assertion but incorrect if it is a quasi-assertion or an assertion for special reasons.⁹

Accepting *P* means that we treat (thoughts or utterances of) *P* as correct or appropriate according to *some* standard of correctness. Treating *P* as correct or appropriate according to some standard of correctness is not just thinking that *there is* a standard of correctness according to which *P* is appropriate. It is to think of a particular practice with *P* — asserting or pseudo-asserting it, believing it, pretending it, etc. — as governed by those standards *qua* whatever practice it is. So knowing that there is a way in which saying ‘Sherlock Holmes was a detective’ is correct *if* taken as fictional is not a matter of accepting it unless one thinks of saying ‘Sherlock Holmes was a detective’ *as* saying something fictional.

The picture of acceptance I am sketching here refers to ‘correctness’, ‘appropriateness’ and ‘standards’ of correctness, all normative notions. The idea (which is not novel) is that we can understand acceptance and its different modes by appealing to the norms which govern particular types of acceptance. But it is also an *internalist* characterisation: it is *not* that acceptance is constitutively tied to the norms which in fact determine correctness (though these are not just accidentally related), but rather that acceptance is a matter of *treating* some claim as subject to a particular norm (so it remains an open question thus far whether the norm a claim is *treated as* subject to is in fact the one that determines correctness objectively).¹⁰

⁹ So note that we are not concerned here with ‘constitutive norms’ such as Williamson (2000: ch.11) appeals to in the analysis of what assertion *is*: my point is about when asserting, for example, is *appropriate*, not when it really counts as assertion. I want to remain agnostic about the constitutive question (with respect to assertion).

¹⁰ This idea is connected to a widely addressed theme, discussed in work of McDowell’s (e.g. McDowell, 1994), of Travis’s (see, e.g., Travis, 2006: ch2, sec.4), and of many others’ (including Frege, Russell and the early Wittgenstein) which is concerned with preserving the possibility, on a given account, of *false judgement*. In this context, of course, the issue is not to do with preserving the possibility of false judgement in the

The thought is that to accept *P* is to treat *P* (i) as subject to some particular correctness condition, *C*, and (ii) as satisfying *C*. (Obviously, ‘correctness condition’ here is *not* supposed to mean ‘truth condition’ in the truth-conditional semantics sense.)¹¹ Nothing ought to be off-putting

sense discussed by those just cited – it is not the possibility of false *beliefs* or judgements in the sense of attitudes towards (or involving) *truth* that I am concerned with. But the connection is that it is, in general, a bad idea to link the characterisation of mental or psychological states in too close a way to the conditions that must be met for being in those states to be *justified*, for in all sorts of cases (including but also besides judgement) we want – indeed, need – to be able to make sense of a person being in that state *wrongly*. Put another way, there are ways of being mistaken or going awry in one’s mental life which do not reduce to false judgement in the narrower sense. This way of thinking of the issue is connected, in turn, to themes in Gibbard’s, 1990, treatment of *choices* and *feelings* as, respectively, *wise* or *apt*.

¹¹ Another explanation of the ‘internalist’ nature of this schema, for those who want it: The ‘external role’ of correctness conditions or norms for claims is to determine correctness. The external role of correctness conditions is to determine when, and, in collaboration with the relevant facts about whether the conditions are *met*, whether, a claim *is* in fact correct. Such a role is external to the attitudes of the person whose acceptance of the claim is at issue, in the sense that their acceptance of the claim (or the claim itself, in a context) is or is not (would be or would not be) correct regardless of whether they *do* accept it. The ‘internal role’ of correctness conditions or norms for claims is to be invoked by a person *in* accepting a claim. That is the role played by correctness conditions in the characterisation of acceptance with which the paragraph to this footnote opens. A person’s acceptance of *P* (in a context) can be correct or incorrect; that is a matter of which correctness conditions objectively determine correctness (i.e. which conditions *in fact* govern correctness), and whether *those* conditions are *in fact* met. The subjective use to which correctness conditions are put in my account of acceptance, though, is to feature as objects of a person’s attitudes *in* accepting *P*: if I treat *P* as subject to *C*, and as satisfying *C*, I count as accepting *P*, *whether or not C is, in fact, the condition which objectively governs the correctness (in the context) of P, and whether or not P in fact satisfies either C or, if different, the condition which in fact governs its correctness in the context.*

This is just to say – or it is just intended to be a way of respecting the thought that – a person can accept things in funny ways, believing, for example, that Sherlock Holmes really did live on Baker Street on the basis of Conan Doyle’s stories. What is going on in such cases, according to my picture of acceptance, is that such a person is treating a claim as subject to an inappropriate criterion of correctness, failing, for example, to see that *qua stories* the Sherlock Holmes tales and the claims made within them do not depend for their acceptability on their *truth*, such that to accept them is not properly to think of the condition being met by them as a truth (or evidence) condition. (Of course, such a person *also* makes the mistake of thinking it true, as *satisfying* that (in fact) inappropriate condition!) Nonetheless, such a mistake is (though a mistake) *possible*, and it helps to illustrate the difference between the role of correctness conditions in characterising how or whether a person accepts a claim, and the *different* role of correctness in determining the objective normative status of person’s accepting a claim. (Note that the same can be said for a person who refuses to accept that Sherlock

about this idea from the perspective of someone eager to avoid an over-intellectualist account of acceptance, for nothing I have said so far (nor anything I go on to rely upon) requires 'treating as' in any particular way. So, for example, it is not required, for all that's been said, that a person explicitly represents to themselves the correctness condition in question or the meeting of that condition. If it is to the reader's taste (it is not to mine, but that is by the bye) to think of 'treating as' subject to a particular correctness condition or as meeting that condition as something like *behaving as if that condition applied and/or were met*, or *behaving as if one thought that the condition applied and/or were met*, then that is (for the purposes of this discussion) just as good as thinking of 'treating as' here in a more explicitly *thought*-involving way.

Also, for our purposes it does not much matter which 'truth-associated' correctness condition is relevant, for the point will be to contrast modes of acceptance (belief) which have *something* to do with truth and those which do not. Note also that throughout I employ the term 'truth' and its cognates for the property of things being the way they are said or thought to be. A more careful presentation of the material here would speak of norms relating to the relationship between what is said or thought and what that which is spoken of or thought about is actually like. The virtue of that more careful presentation would be that it would respect the thought that conceptions of *truth* are sophisticated meta-semantic commitments, commitments which a person might very well

Holmes lived on Baker Street, even for the purposes of literary discussion, because they do not think that claim satisfies a truth-associated condition: their mistake is not to think that is it not true that Holmes lived on Baker Street (he did not, it's just a story, so they are quite right about *that*); it is to think that *that* fact is appropriate grounds on which to refuse to accept the claim (in the context of literary discussion, at least), for they have mistakenly taken the truth-associated correctness condition to be the one to which the claim is, in the context, subject. So rejection can be used just as well as acceptance to mark the difference internal and external roles of correctness conditions in relation to acceptance.)

lack and yet accept all sorts of claims which, on the model I have presented as stated and taken literally, they could not since they would lack a conception of *truth* required for playing the role of perceived correctness condition (or of grounding the truth-associated correctness condition). I employ 'truth' in the less formal way, though, with what I take to be reasonable endorsement from standard usage. And nothing in what I have said here about the limitations imposed by taking my talk of 'truth' as referring to something of the sort formal semanticists and meta-semanticists are interested in harms my point, for everything I say about 'truth' can, if the reader prefers, be translated into talk of that the property of things being the way they are said or thought to be.

(While we are noting aspects of the proposal, it is an interesting question whether the theory of acceptance proposed here requires abandoning the traditional belief/desire model of action. It seems to me that it does, though I have nothing to offer here about the ramifications of that. Suffice it to say (with, for example, Bratman, 1987) that if an adequate theory of some mental/psychological states requires treating those states as neither beliefs nor desires but as nonetheless apt to play the sort of role usually reserved in action theory for beliefs or desires, then so be it. I hope to be motivating the thought that an adequate account of acceptance *does* require a notion of acceptance that is neither belief nor desire; and it seems obvious to me – as I hope it will to the reader, especially after reading Chapters Four and Five – that acceptance can be just like belief in its action-related profile in at least some circumstances.)

We have, I think, a feature of all and only modes of acceptance. Think of any proposition that you accept, *P*. Is there some way in which thinking *P*, or saying *P*, seems correct? I doubt there is any *P* which you would be prepared to say you accept for which no such way of seeming

correct could be found. Of course this feature of acceptance allows for just the sort of pluralism that the fictionalist wants. There is *some* way in which it is correct or appropriate to say or think what is true. There is *another* way in which it is correct to say or think what is useful, or practically necessary. And yet *other* ways in which it is correct to say particular things in a play, to do particular things in a game, to say things about Sherlock Holmes and so on. These are not the same ways. All the same, modes of acceptance appropriate to each of these reasons for saying or thinking things have the feature of making it seem correct or appropriate to say or think them.

It is plausible, I think, that other features of acceptance are 'theorems' of this feature — they derive from it. For example, Gibbard's 'responsiveness to the demand for consistency' which marks acceptance out from mere internalization (which is also a matter of 'tending to avow' norms) can be thought of as what it is necessary to manifest in one's thinking if one is to see some norm as correct or appropriate: if one were to fail to be responsive to the demand that one's normative commitments be stable, what could the thought that those normative commitments are *correct* consist in? Similarly, Kalderon (2005a: pp3-8) suggests that 'full acceptance ends inquiry'. This feature of acceptance can be thought of as following from the feature I described, also: once a person sees a claim as *correct* or *appropriate*, they have good reason to enquire no further, for whatever reasons they have for thinking that their claim is correct or appropriate will do quite well for the purpose of satisfying whatever doubt there might be about the propriety of the claim which prompted enquiry in the first place. Put another way, enquiry is, most generally, enquiry into correctness or appropriateness, so coming to see the correctness or appropriateness of a claim is bound to be sufficient reason to end enquiry. Of course it is not *necessary* to end enquiry once the

correctness — by *some* standard — of a claim has been established: we might very well decide that there are other measures of correctness against which we wish to measure our claim. But it is obvious that *no* consideration is sufficient to end *all* inquiry in this broad sense, whatever our theory of acceptance.

What further feature of *belief* distinguishes it from other modes of acceptance? Believing *P* involves thinking of *P* as the correct or appropriate thing to think or say in various particular contexts; having this general feature is what makes belief a mode of acceptance. What is distinctive about belief is that those particular contexts are ones to do with what things are really like. To believe *P* is to adopt an attitude towards *P*'s *truth*. It is to see the correctness of thinking that *p* as turning on whether it is *true* that *p*. It is, in short, to be responsive to a truth- (or truth-associated-) norm of belief.¹² Other modes of accepting *P* do not involve adopting an attitude to the truth of *P* at all. When we accept that Sherlock Holmes was a detective *qua* correct fictional claim, we do not think that it is literally *true*, and we do not think of the truth or falsity of the claim as having any relevance to the fact that saying 'Sherlock Holmes was a detective' is more correct or appropriate than saying 'Sherlock Holmes was a Rastafarian'.

Nondoxastic acceptance often goes hand in hand with belief: most people accept that Sherlock Holmes was a detective and reject that he was a Rastafarian in the way I described, but also *believe* that *according to the Conan Doyle stories* Sherlock Holmes was a detective and not a Rastafarian. Their belief *is* truth-normed, since it is an attitude towards what a particular story *actually* says. But 'Sherlock Holmes was a

¹² It is, of course, contentious quite what the truth-norm or truth-associated-norm of belief *is*. And for all I say here, there is no reason to think that truth is the *only* norm of belief. (See n2, above.)

detective' and 'according to the Conan Doyle stories, Sherlock Holmes was a detective' express different propositions (if they were identical, there would be something presciently post-modern about the Holmes novels, since every Sherlock Holmes claim would in fact be a reference to the work itself! See Joyce, 2005 on the difference between describing a story and telling it), so we ought not to be tempted to say that acceptance of the Sherlock Holmes claim *just is* belief in the Conan Doyle story claim. Believing the Sherlock Holmes claim would amount to thinking of the correctness of that claim as turning on whether the claim itself were true, not on whether some *other* (e.g. Conan Doyle story) claim were true. So even though belief and mere acceptance often travel in close company, they are not the same nor is either strictly reducible to the other.¹³

Notice also how this picture differentiates acceptance from assumption in a plausible way. In *assuming P* (e.g. for the purposes of *reductio*), I do not come to see *P* as correct — indeed, as noted in section 2, above, it might be my conviction that *P* is *incorrect* that explains why I go in for assuming it at all. I may see it as correct *to assume P*, but that is not the same as seeing *P* as correct. (For more on this broad theme, see section 5, below.)

We now have a way of (i) distinguishing acceptance from other attitudes, (ii) distinguishing modes of acceptance and (iii) saying what is distinctive about belief, as a mode of acceptance. This is just what the fictionalist needs.

4. Horwich's Objection

¹³ This is important for a fictionalist such as van Fraassen, 1980, who speaks a lot of, e.g., empirical adequacy and the propriety of acceptance of some claim turning on beliefs about that claim's empirical adequacy – beliefs *about empirical adequacy of P* are not beliefs *that p*.

Paul Horwich doubts that it is possible to distinguish belief from what I want to call nondoxastic acceptance. His core thought, directed here against 'instrumentalists' in philosophy of science (whom Horwich does not, unfortunately, seem to distinguish from fictionalists) seems to be that:

If we tried to formulate a psychological theory of the nature of belief, it would be plausible to treat beliefs as states with a particular kind of causal role. This would consist in such features as generating certain predictions, prompting certain utterances, being caused by certain observations, entering in characteristic ways into inferential relations, playing a certain part in deliberation, and so on. But that is to define belief in exactly the way instrumentalists characterize acceptance. ([1991]: p89)

The point I want to raise against Horwich here is that, read one way, even Horwich's own chosen marks of belief *can* be made to differ from marks of nondoxastic acceptance. It does not contradict anything I have said about acceptance in general, above, (which, as we saw, allows for just the sort of differentiation the fictionalist wants) to say that belief has the feature of, for example, 'prompting certain utterances' — after all, on the account I sketched above, *believing P* is likely to prompt (in relevant contexts) utterances of '*P* is true', or 'no, you're wrong, *P* not not-*P*!' etc. Belief, that is, is likely to prompt utterances about what things are really like, for belief has the feature of being to do with taking an attitude towards the *truth* of *P* (treating *P* as correct in virtue of its truth). If Horwich is right, the utterances which believing *P* is likely to prompt are the same as the utterances which (nodoxastically) 'accepting' *P* is likely to prompt; this is what grounds his charge against the 'instrumentalist' (fictionalist) that they have tried to introduce a distinction without a difference. But surely it is plausible that nondoxastically accepting *P* can be likely to prompt utterances which believing *P* would *not* be likely to prompt: we have been using a claim about Sherlock Holmes to make just

this sort of point: nondoxastically accepting that Sherlock Holmes was a detective is not likely to prompt us to say 'It's *true* that Sherlock Holmes was a detective'; nor, for that matter, is it likely to prompt any of the sorts of *behaviours* which depend for their rationale on thinking that there really is a Sherlock Holmes (such as looking for evidence of him in the historical record).

Perhaps it is harder to find examples in *science* where what we would in fact, be likely to predict, utter etc. depends upon whether we believe or 'merely accept' *P*. But that *might* just be because scientists and those thinking about science have got into a particular habit: the habit of confusing an enquiry whose *object* is what the world is really like with the *norms* governing that enquiry. It is not a necessary truth that practices of saying things about how the world really is are *normatively governed* by the *truth* of those claims. That is why we can think of the Sherlock Holmes claim as a claim about a real person but as a *false* claim about a real person *the falsity of which does not debar correctness*. But if, as it happens, nobody thinks of science in a similar way, then people will be likely to more or less automatically behave as if *P* is true for some scientific *P* whenever they accept *P*, nondoxastically or not, for it will seem to them that since *P* is *about* how things really are it must follow that (for example) uttering '*P* is *true*' must be *appropriate* in virtue of the sort of claim *P* is. That is, the explanation for whatever similarity there might be in the utterances and behaviour of those who nondoxastically accept and those who believe some scientific claims *might* be that those whose utterances and behaviour are appealed to *misunderstand* their own commitments in respect of what utterances and behaviour *really make sense* given those commitments, which is to say given the way in which they accept the claims in question.

If the fictionalist is right then the habit of straightforwardly associating the *object* of some discourse with the *norms governing* that discourse is a mistake. And if the fact that acceptance of some scientific *P* is likely to prompt utterances associated with *belief* that *P* reflects an unthinking *mistake* on the part of those whose propensity to utter is being appealed to, there is little reason to treat that evidence against the fictionalist's distinction as admissible (or at least it is certainly not persuasive).¹⁴

Elsewhere, Horwich makes essentially the same point in the terminology of the 'belief box' psychology of Fodor and of Shiffer. For something (a proposition or sentence) to be a person's belief is (roughly) for it to be such that (a) sensory experiences cause associated sentences to also go into the belief box, (b) it 'bring[s] it about, in virtue of inference rules, that certain other sentences [or propositions] are also there', (c) 'some such sentences are relied upon in practical reasoning' and (d) 'when (and only when) a sentence is in the box there is a disposition to utter it' (Horwich, 2006: pp192-3).

Straight away, it is obvious that the deck is stacked against the fictionalist, or anyone who wants to distinguish belief from nondoxastic acceptance, here: if for some proposition to be a belief for someone is *inter alia* for that person to be disposed to utter a sentence expressing it when *and only when* that proposition (or a sentence expressing it) is in *that* box, then it follows that any mode of acceptance of which it is a feature that a person is disposed to utter (sentences expressing) what is accepted *must* be belief. This deck-stacking, however, doesn't seem remotely plausible

¹⁴ Horwich's way of thinking about acceptance/belief appeals to the sorts of dispositions I said were unnecessarily problematic or controversial. But my objection to Horwich does not turn on that feature of his view. Rather, the objection is that *even granting his assumptions about what would determine whether an attitude counts as acceptance*, it is possible to distinguish modes of acceptance.

anyway, on common-sense grounds entirely independent of fictionalist motives, for it is surely not plausible, for example, that a person who is disposed to tell a particular lie – surely a case of uttering a sentence expressing a proposition – thereby counts as *believing* the lie, as having the propositional content of the lie in their belief box. Presumably conditions (a)-(d) are intended to be *jointly* sufficient for a proposition being in a person's belief box; but then (d) ought to omit the '*only* when' direction, for what (d) is supposed to add to (a)-(c) is, presumably, that even if (a)-(c) are satisfied a person does not believe *P* if they are not at all disposed (in the relevant, yet to be specified sense) to utter sentences expressing it, *not* that a person would *only* be so disposed (for being disposed to lie is an all too familiar datum).

Horwich then, more or less correctly, notes:

The fictionalist (*et al.*) would have to suppose, in effect, that there is *another* box for sentences – let's call it the fiction box – and that someone can choose, when convinced by skeptical arguments directed at a certain discourse, to shift its sentences from his belief box to his fiction box. He will then qualify as 'accepting' the propositions they express but, allegedly, no longer as believing them. (2006: p193)

What, though, would be wrong with that? According to Horwich:

[T]he problem is that, if there is to be no revision of the practice, any such sentence will now have to go into the fiction box in whatever circumstances it would have previously gone into the belief box. In addition, the consequences of a sentence being in the fiction box vis-à-vis inference, utterance, and action will have to be the same as what the consequences would have been of its being in the belief box. Therefore, even though someone's saying "p" will, allegedly, no longer express his *belief* that, but merely his *acceptance* of that proposition, this won't have any causal significance; for the psychological import of the fiction box is just the same as that of the belief box. So there would appear to be no difference at all between these supposedly different boxes – no difference at all between believing *that p* and the attitude of *accepting that p*, which is being urged as a less committal alternative. Thus it is an illusion that we

might persevere in deploying sentences exactly as if we believed what they literally express, but without really doing so. (2006: p193)

Let us accept that the challenge of showing that there are non-doxastic modes of acceptance does indeed amount to the challenge of saying what the difference, in terms of 'inference, utterance and action', between being in the belief box and being in the fiction box amounts to. In response to Horwich's earlier formulation of his criticism, I have already said, *pace* Horwich, that I think there are grounds to doubt that the *consequences* of being a belief and being non-doxastically accepted are identical: we would not be inclined to *say that P is true* (in contexts where talk of truth is to be taken seriously) if we did not believe, but non-doxastically accepted, it, nor to behave in ways the rationales of which depend upon thinking of *P* as true.

Horwich is not sufficiently sensitive to the distinction we ought to draw between practices. (See the previous chapter.) We ought to distinguish, obviously, between mathematics and metaphysics, for example, not just in terms of their *objects* but in terms of what it is to *do* mathematical enquiry and metaphysical enquiry. Part of the difference between doing mathematics and doing metaphysics is, of course, that we are asking different questions, often about different things. But that is not always so. Faced with the proposition that there is a prime number between six and eight, a mathematician may ask 'Is there a prime between six and eight?'; and so might a metaphysician. When a mathematician asks that question, they are likely to be taking for granted that *some* number exists between six and eight, and that the numbers six and eight exist, but wondering whether it is one of the *prime* ones or not. On the other hand, when a metaphysician asks that question, they are not really concerned with whether the number in question is *prime*; they want to know whether there is a number there *at all* (presumably because they

want to know whether there are *any* numbers at all). So it is not just the things talked about which differentiate practices, but the interests taken in them, the standards of evaluation appropriate to claims about them, and the *point* of talking about them, too.

So it is quite wrong of Horwich to infer, as he does, from the fact that the hermeneutic fictionalist wants to 'leave everything as it is' with respect to the practice in question (theoretical physics, mathematics, modal discourse, moral enquiry or whatever) to the conclusion that there is bound to be *no* difference in terms of 'inference, utterance and action' between being in the belief box and being in the fiction box. True, leaving everything as it is with respect to the target discourse may well mean denying any such difference *with respect to that discourse*. But that does not mean no difference *at all*: it might be that non-doxastically accepting some ontologically committing mathematical claims, for example, involves no relevant difference from believing them *for the purposes of mathematics*, or for a person *qua mathematician*; but that does not mean that none of a person's inferences, utterances or actions *outside of that practice* will be materially different. For example, their utterances *qua mathematician* might be just the same (they would still say 'There exists a prime between six and eight') whether their acceptance were doxastic or non-doxastic. But *qua metaphysician*, their utterances would most certainly *not* be the same: *qua metaphysician*, they would not be willing to say 'There exists a prime between six and eight' if, as Horwich said, 'when convinced by skeptical arguments directed at a certain discourse, [he] shift[ed] its sentences from his belief box to his fiction box'.

The only way to block this defence of non-doxastic acceptance against Horwich's argument would seem to be to insist that propositional attitude states such as belief are determined by reference to some privileged practice(s): if it were the case that the conditions on counting

as belonging in the belief box being met *with respect to some particular practice(s)* were sufficient for belief *tout court*, and if the fictionalist's target practice(s) were some such privileged practice(s), then it would be enough for belief *tout court* if there were no Horwich-type differences *just within the target practice with the discourse*.¹⁵ This, of course, is equivalent to Horwich's apparent assumption that the target practice is the *only* practice of interest.

It is plausible that there is a class of privileged practices with respect to which the nature of a person's acceptance would be determined, if it were to be determined in the Horwich type way at all. After all, we surely only want to look at the inferences, utterances and actions a person undertakes in the course of some sort of serious enquiry, deliberation or suchlike. But there is no good reason to expect the fictionalist's target practices to be *those and only those* practices of serious enquiry etc. For a start, if metaphysics is a practice of serious enquiry, then the example I gave above shows that even within serious enquiry there can be differences of just the sort Horwich requires and claims not to be able to find. Any more fine-grained specification of which practices, *within* serious enquiry, are to count as determining states of acceptance (ruling in mathematical enquiry, for example, but ruling out metaphysics) would seem even more *ad hoc* than the doxastic/non-doxastic distinction is supposed to be.¹⁶

¹⁵ Actually, even this wouldn't be quite right: it would have to be that the target practice(s) *exhaust* the class of privileged practices.

¹⁶ It is, I think, possible to settle this debate (to at least the required level of satisfaction) without addressing the issue of Horwich's minimalist theory of truth (Horwich, 1998). It might be tempting to think that on a theory such as Horwich's, there is no respectable sense of 'true' robust enough to allow for any difference between uttering '*P*' and uttering '*P* is true', so that the sort of difference in dispositions to utter that I made something of are not available. However, Horwich is not in fact committed to denying that there is a truth property picked out by a semantically respectable truth predicate (indeed he explicitly rejects the sorts of theories which deny such things: Horwich, 1998: pp38-40), so it perfectly in order to press Horwich on apparent differences in utterance

5. Fictionalism and Pragmatic Norms

For the fictionalist, it is appropriate for mathematicians, for example, to go on with their literally ontologically committing mathematical claims without *believing* them (for believing them, since they are literally ontologically committing, would commit mathematicians to some ontological undesirables such as sets, for example). Just as engaging with the Sherlock Holmes fiction involves some sort of acceptance other than belief — for it is correct to say and think that Sherlock Holmes was a detective, not to just *reject* that claim outright — it is part of *doing mathematics* (or of doing science which involves mathematics) to accept — but, of course, not necessarily to believe — the ontologically committing claims: mathematicians cannot afford to *reject* their mathematical claims just because they don't believe them, so some other form of acceptance must be appropriate. (See, for example, Field, 1980; 1989.) And just as in

dispositions in respect of our tending to apply or not apply the truth predicate. Perhaps it would be objected that although Horwich countenances such a predicate he denies that there is anything of substance to said in analysing it, beyond saying that it refers to a property, the analysis of which is also entirely vacuous, so that whilst we might be disposed to apply or not apply the predicate in various contexts, that does not amount to a serious difference in what is said. That does not seem like a strong defence, though, for once it is granted that the truth predicate is semantically respectable (whatever the interest or otherwise of the property to which it refers), there is every reason to be interested in the difference between a person's disposition to employ it and their disposition to not. Of course many deflationists (following Frege, [1918]) are prepared to say that the content of '*P*' and of '*P* is true' is identical, and that there *is no* semantic difference upon which different utterance dispositions can be based. All I shall say about that here – which goes for Horwich's theory, too, if it turns out that some resource provided by his theory helps to resist my criticisms after all – is that minimalism or deflationism about truth is a high price to pay for an adequate theory of acceptance, and that unless we are already persuaded on entirely independent grounds that such a theory of truth is correct we have no reason to appeal to its resources in deciding whether to distinguish nondoxastic acceptance from belief. And anyway, *utterances* of '*P* is true' were only one type of difference I adverted to; it is unclear whether, e.g., *reliance on* some claim in different (not-target) contexts could be treated by a minimalist theory of truth in such a way as to undermine the point I want to make.

the fiction case, mere acceptance and belief may travel in close company: merely accepting some mathematical or scientific claims might be closely tied to *believing* that those claims are useful or ‘empirically adequate’ (see van Fraassen, 1980).

Now that we have a way of distinguishing belief from nondoxastic acceptance (a way which is independently plausible), we can state the fictionalist’s proposal for some target discourse, *S*, in terms of the more precise semantic and psychological theses I opened with. The fictionalist thinks that the claims of *S* are truth-apt and purport to describe facts (this is their representationalism), but that whilst they are truth-*apt* there is some practice involving *S* in which acceptance is not truth-*normed*. For that practice, it is appropriate to accept claims of *S* since there are reasons to think of those claims as *correct* or *appropriate* for some end. But since that correctness or appropriateness is not determined by the literal *truth* of the claims, the acceptance that the fictionalist diagnoses or recommends is *nondoxastic* acceptance.

Why does the fictionalist insist upon the correctness or appropriateness of the claims of *S*? The important aspect of the fictionalist’s view here is that there is reason to think of the claims of *S* as correct or appropriate *for some particular practice* with *S* (as we saw in the previous chapter). To take Field’s (1980; 1989) mathematical fictionalism again, there are pragmatic reasons for *mathematicians* and scientists who employ mathematics to regard all sorts of mathematical results — including those which are ontologically committing — as correct. It will not do to just treat all mathematical results as *assumptions*, for we can just as well assume *P* or its negation whereas it matters in mathematics whether we accept *P* or accept not-*P*. (Compare: we do not *assume* that Sherlock Holmes was a detective and not a Rastafarian, we *accept* it.) Field thinks it would be problematic if accepting mathematical claims

had to amount to *believing* them, since they are — literally construed as quantifying over sets etc. — *false* and ought not to be believed. It would be no skin off *philosophers* noses (*qua* philosophers, at least) not to accept mathematical claims any more, for philosophers (*qua* worried about the ontology of maths, at least) are interested in what to believe, in what is true. Mathematicians, though, would be up a gum tree if acceptance of mathematical claims had to be truth-normed in the way belief is. So they had better *nondoxastically accept* those claims, to accept then in such a way that their treating them as correct depends upon some condition of correctness other than truth. Likewise for van Fraassen's, 1980 'constructive empiricism' for scientists, or Rosen's, 1990 suggestion that possible world semantics make use of the posits of modal realism without being committed to their truth. In short, the fictionalist typically insists upon the correctness or appropriateness (in spite of the falsity) of the claims of *S* for some practice for *pragmatic* reasons.

These pragmatic reasons for insisting on the correctness or appropriateness of the claims of *S* (for some practice) get into the acceptance norms for that practice (with respect to *S*) in virtue of acceptance being that attitude which has the feature of seeing some claim as correct in some way. But we should mark a distinction between *having a use for* some claims and *accepting* them. The distinction is between saying something about acceptance norms and saying things about the much broader notion of reasons for engaging with some body of claims. To see what this distinction amounts to, compare the suggestion (by Wolff, 2002: p101) that 'we value the work of the greatest philosophers for their power, rigour, depth, inventiveness, insight, originality, systematic vision, and, no doubt, other virtues too' and that '[t]ruth, or at least the whole truth and nothing but the truth, seems way down the list'. This suggestion about the reasons we have for reading and thinking

about the work of the great philosophers is *not* the suggestion that we *accept* the claims of those philosophers, but that we have reasons to be aware of them and to think carefully about them. These reasons might themselves be philosophical reasons: seeing where a starting point leads when carefully worked out, for example, can be philosophically important. Or the reasons might be of a more general or historical nature: it is good to know what influential people have thought, and why. But these sorts of reasons do not generate reasons to think of the claims of the great philosophers as *correct*. There are good philosophical and historical motives for reading Marx on the nature of history. But it would be silly to say that I think Marx was *correct* to say what he did about history *because* of those philosophical and historical motives. (The requirement to read philosophers charitably does not, of course, amount to any demand to think that they are *right*, nor to think that their remarks are appropriate in any other sense!) The practices in which the fictionalist is interested, though, are not ones for which such agnosticism about correctness is viable. It is not just that mathematicians or scientists, for example, have non-truth-directed reasons to *engage with* the claims of *S*; it is that they have good pragmatic reasons to regard some particular claims of *S* as *correct*, for it matters to their calculations, investigations and so on which claims to rely on and which to jettison. So the fictionalist's proposal is one for which nothing less than pragmatic *correctness norms* will do.

6. Not Just Practicalities

Not all varieties of fictionalism appeal to *pragmatic* norms though (as we noted in the previous chapter). Kalderon, 2005a, for example, develops a fictionalist suggestion which replaces belief in moral propositions with what I am calling nondoxastic acceptance of those propositions, but does

not argue for any distinctively *pragmatic* norms with which to replace the truth-norm of belief.¹⁷ He argues that:

In accepting a moral sentence that he understands, a competent speaker reconfigures his affective sensibility so as to render salient, in a phenomenologically vivid manner, the moral reasons apparently available in the circumstance, as he understands it. In accepting a moral sentence that he understands, a competent speaker quite literally decides how he feels about things. It is the structure of a person's moral consciousness, and not some further fact, that constitutes the relevant kind of affect. (2005a: pp50-1)

Such an 'affect-centred' way of characterizing a person's acceptance is clearly not the same sort of thought about acceptance as the pragmatic conception of the mathematical and scientific fictionalists (and others).

But the theory of acceptance sketched above does just as well with respect to affect-centred as to pragmatic models. For it is clear from what Kalderon writes that in accepting some moral claim, a person is in precisely the state the theory predicts: in virtue of the phenomenal character of the situation they are considering, including the features of that situation (including the moral reasons) which seem salient to them and how it therefore feels to them, they will be in a position to see the moral claim as appropriate. In seeing something as cruel, for example, — in being presented with a case certain features of which strike them as overwhelming reasons to disapprove — a person will see utterances or thoughts of 'you *mustn't* do that, it's *cruel*' as entirely appropriate. Indeed, it will seem to them that it is entirely *inappropriate* to think anything else about the case. Note that this near-truism does not commit us to the thought that a person in such a situation would have to see the correctness of the moral claim as turning on its *truth* — the push to see the claim as appropriate comes from the way things strike them, and they

¹⁷ See Joyce, 2001; 2005 for a form of moral fictionalism that employs a pragmatic norm in place of the truth-norm.

need not necessarily be committed to thinking that the way things strike them accurately represents any extra-phenomenal reality, though of course they *might* be committed to such a view.¹⁸ Thus the feature of acceptance I drew attention to in section 3 is a feature of the sort of acceptance affect-centred theorists have in mind just as it is for the pragmatic theorists (and, I suggest, for any other plausible notion of acceptance).

7. Conclusion

Regardless of whether or not we want to be fictionalists, there is good reason to recognise various modes of acceptance, individuated by the various correctness or appropriateness conditions a claim is treated as subject to, in the context of some practice or activity, and as satisfying. Those, such as Horwich, who doubt that belief differs from nondoxastic acceptance on the grounds that there is no relevant difference to be found between them must have an unorthodox (and seemingly unmotivated) idea of what sorts of difference would count as relevant.

I have said that fictionalists are committed to semantic representationalism: they think that the claims of *S* purport to say how things really are. Those claims are, therefore, truth-apt. I also said that fictionalists mark a distinction between *believing P* and *nondoxastically*

¹⁸ Of course the suggestion that the relevant notion of truth here is to do with 'extra-phenomenal reality' is going to strike some as a contentious (perhaps naive) oversimplification. Quasi-realist minimalists such as Blackburn (e.g. Blackburn, 1984: ch6; 1993) and pragmatists such as Putnam (e.g. Putnam, 2004) certainly won't think it obvious that moral truths are bound, *qua* truths, to be independent of us in all ways, and would (I think) be quite content to think of the person as being committed to the *truth* of the claim in virtue of its correctness. As, interestingly, would Stevenson (1963b: sec.8) who came to object vociferously to what he regarded as a misreading of his emotivism according to which moral claims are neither true nor false: Stevenson actually adopted Ramsey's deflationism about truth and regarded (with some qualification) the truth of moral claims as tantamount to their agreeableness.

accepting P, and that they are committed to nondoxasticism: they think that appropriate acceptance of the claims of *S*, for some practice with *S*, is some mode of nondoxastic acceptance. And I offered a way of understanding acceptance and the distinctions between modes of acceptance in terms of putting one in a position to see thinking or uttering *P* as correct or appropriate, *qua* whichever sort of thought or utterance it is, in the hope of starting to motivate and explain the distinction upon which the fictionalist relies.

Most fictionalists (*all* committed fictionalists currently in print, to my knowledge) are non-realists (they doubt that there really are mathematical abstracta, natural or non-natural moral properties, scientific unobservables, real possible worlds etc., or at least that we can confidently assert that there are). A non-realist, faced with the desirability of not dispensing with *correctness* (as discussed in section 5, above) has good reason to resort to fictionalism. But what if you are a realist who thinks that the claims of *S* are true? You might, for example, be some sort of realist about fictional characters (see discussion in Sainsbury, 2009), but think that whilst there happens to be a real (non-actual, or non-concrete) Sherlock Holmes who really has (maybe *encodes*: see, for example, Zalta, 1988) the property of being a detective, *acceptance* of that claim — at least for those who are just interested in the stories, rather than philosophy of fiction — is not governed by whether or not that fact obtains: it would have been just as *correct* to say that Sherlock Holmes was a detective if that fact *hadn't* obtained, if fictional realism (realism about fictions) had turned out to be false. Realists about some discourse can avail themselves of just the same distinction between correctness in virtue of truth and correctness in virtue of the meeting of some other condition as non-realists can, and can even go so far as to avail themselves of *nondoxasticism*.

In Part II, I will explore *why* a theorist of some practice – ordinary moral judgement, deliberation and commitment, specifically – might be attracted to nondoxasticism, and will suggest that these are reasons which a realist might have. Assuming that realists are committed to semantic representationalism more or less by default, adopting nondoxasticism for some practice with *S* amounts to fictionalism about that practice with *S*. Thus, Part II will begin to *motivate* realistic fictionalism.

Part II

The Nondoxastic Character of Moral Commitment

‘That is honourable, I mean to say, it’s humane! You wanted to avoid gratitude, I saw! And although I cannot, I confess, in principle sympathise with private charity, for it not only fails to eradicate the evil but even promotes it, yet I must admit that I saw your action with pleasure – yes, yes, I liked it.’

Lebeziátnikov, in *Crime and Punishment*¹

1. Introduction

In Part I we saw in very general terms what a fictionalist theory of some target discourse might be like, and we saw that nothing at that general level stands in the way of what I am calling realistic fictionalism. Non-realists might profitably avail themselves of fictionalism, as some have. But realists could also avail themselves of fictionalism and whatever benefits it brings. The question, then, is whether a realist about some discourse has any good *reason* to thus avail themselves. It would be of merely taxonomical interest if realistic fictionalism were a position in logical space which nobody has any reason for occupying, if fictionalism brought benefits only for the non-realist. So now, in Part II, I go some way towards addressing that question, by suggesting that the best theory of our *ordinary moral judgements, deliberation and commitments* might well be realistic fictionalism.

The relation between our moral judgements, moral deliberation and moral commitment is, no doubt, complex and interesting. In this thesis, though, I treat moral judgement, deliberation and commitment, and the relations between them, in a relatively straightforward way: I take it that our

¹ Dostoevsky ([1866]: p317). For some critical discussion of this passage, linking it to the point of this chapter, see the excursus at the end of this chapter.

moral *commitments* are those claims or principles we accept in a relatively stable way (what we might call 'passing moral thoughts' are not commitments in my sense), and I am open to the idea that our moral commitments (like our other commitments) *tend* to play a role in explanations of our actions more regularly and/or reliably than thoughts in general do (which is *not* to say that it is part of what it *is* to have a (moral) commitment that we tend to act on it); I take it that moral *judgements* are in many ways the occurrent analogues of our commitments, which (I suppose) are state-like (in just the way that judgement is naturally taken to be the occurrent analogue of belief, which is state-like), and that we can and often do come to have moral commitments on the basis of making moral judgements, and that we often employ our moral commitments in coming to form moral judgements; and I take it that ordinary moral *deliberation* is a process (or, more realistically, a family of processes) of coming to form moral judgements on the basis of our moral commitments, other moral judgements and all sorts of non-moral commitments and judgements (and, indeed, whatever other sorts of things are supposed to be involved in deliberation generally). So, I take it that a theory of our ordinary moral thought will naturally and necessarily have something to say about our moral commitment, judgement and deliberation, and that such a theory will treat each as at least in part dependent on the other(s): a theory of moral deliberation will mention moral judgements and moral commitments, an understanding of which will naturally involve an understanding of their roles in deliberation.

I intend, here, to do very little justice to the debate which might very well be engendered by my taking a view of moral commitment, judgement and deliberation which treats them as so interdependent; suffice it, for my purposes at least, to assume that a theory of the character of, for example, moral commitment has implications for the theory of moral deliberation: if,

for example, we are convinced that our ordinary moral commitments are nondoxastic, then we have reason to think that the norms governing ordinary moral deliberation must be norms other than those concerned with, for example, truth-preservation. Similarly, if our ordinary moral commitments are typically nondoxastic, it will be natural (given the supposed links between commitment and judgement) to treat our ordinary moral judgements as typically nondoxastic. I do not suppose that a nondoxastic theory of moral judgment and/or commitment is *entailed* by a nondoxastic theory of our ordinary moral commitments, nor that the reverse *entailment* holds. I can readily imagine there being a theory of ordinary moral thought according to which, for example, our moral commitments are typically nondoxastic but the judgements which such commitments often enable and from which such nondoxastic commitments often arise are straightforward beliefs. But it seems to me that such a theory would have to be more complicated and contentious than a theory which treats our ordinary moral commitments and judgements similarly as nondoxastic, and that therefore if a uniformly nondoxastic theory can be motivated and defended then it is, for all the *prima facie* evidence at least, to be preferred. In the remainder of this section, then, I proceed under the assumption that a nondoxastic theory of our ordinary moral commitment underwrites or at least mutually supports nondoxastic theories of ordinary moral judgement and deliberation.

In this chapter I shall begin to develop a positive characterisation of ordinary moral commitment. It will be a fictionalist characterisation, on the now familiar understanding of fictionalism according to which a fictionalist is someone for whom the semantics of their target claims (moral claims in our case) is *representational* whilst the mode of acceptance appropriate to those claims (for some practice, in our case ordinary moral commitment) is *nondoxastic*.

The variety of fictionalism I want to propose does not involve anything like treating moral claims as ‘elements of the fiction of morality’ and acceptance of them as acceptance of a fictional claim, or anything similarly dependent on an understanding of some mode of acceptance involved in accepting non-moral claims about, for example, Sherlock Holmes. (See Chapter One) For this reason, my account is not going to be subject to many of the worries of the sort raised by Stanley, 2001 to hermeneutic forms of fictionalism, such as the fiction- or pretence-assuming form of the ‘Autism Challenge’.² Nor will my account fit neatly into treatments of hermeneutic fictionalism such as Eklund, 2005 which acknowledge the possibility of just two options for the fictionalist, an account based on treating the content of the target discourse as fictional or an account based on treating the mode of acceptance of the target discourse as the mode of acceptance appropriate to accepting fictions.

I shall sketch the model I have in mind, and then say some things which I hope motivate the thought that this model is a good model of our ordinary moral commitment as it often is. In the course of motivating that thought, the model itself will hopefully become clearer. My primary purpose here is to begin to develop the model. I argue elsewhere that there is reason to *want* our moral commitments to be as the model says they are. But a precursor to assessing that sort of claim is to know what the features of the model *are*, so here I set out to provide just that, and to the extent that I attempt to motivate the thought that our ordinary moral commitment is

² Such challenges take the fictionalist’s supposed model of commitment to some sorts of claims as akin to our acceptance of fictional claims, and press the objection that our acceptance of the target claims is not, in fact, akin to our acceptance of fictional claims, or our practices involving some sort of pretence, as supposedly evidenced by the way in which, for example, those typically unable to deal in the usual way with fictions or pretence are perfectly able to deal with the target claims in seemingly the same way as the rest of us. Obviously, whatever you think of the empirical claim upon which such objections rely, no form of fictionalism *not* committed to treating our target commitments as fictional or pretend could be subject to such an objection.

often like that, it is partly to motivate the weaker thought that it *might plausibly* be that way (for if the model is coherent but not psychologically plausible there is no point in *wanting* things to be that way).

2. The Character of Moral Commitment

We shall work within the theory of acceptance developed in Chapter Three. Acceptance of some claim, *P*, amounts to considering *P* correct, as the thing to say or think, according to some norm of correctness. Put another way, accepting *P* is treating *P* (i) as subject to a particular standard of correctness, *C*, and (ii) as satisfying *that* standard of correctness.

What, in this framework of thinking about acceptance – a framework in which it makes sense, indeed it is necessary, to admit modes of acceptance other than belief, where belief is defined as the treating of a claim as subject to and satisfying a *truth*-associated norm governing correctness – is the right way of characterising acceptance of moral claims in the course of ordinary moral deliberation? The picture I want to present is one on which in accepting a moral claim (for the purposes of this chapter and in the interests of brevity I'll omit 'in the course of ordinary moral deliberation' unless required for clarity by context) we typically treat that claim (*qua* moral claim) as subject to a norm of correctness invoking neither truth nor any closely associated notion (such as evidence) but rather notions associated with *being a particular sort of person*.

The proposal is that in thinking something to be, for example, *cruel* and so/or *wrong* a person is typically committed to treating the thought that it is cruel or wrong as being the correct thought to have in virtue of that thought meeting a standard of correctness, seen as particularly appropriate to the circumstance of judging – a standard according to which a thought about the cruelty of that particular act (or whatever) is correct if a person for

whom notions of cruelty and wrongness are significant in their thought and for their practical deliberations would reasonably be committed to that thought.³ And what makes *that* standard seem particularly appropriate to the circumstance of judging is, roughly, that a person typically takes an interest in the plight of others such as to make *being the sort of person who thinks in terms of cruelty* something that is valued. Or, again, in judging something (an action or possible action, presumably) to be wrong for its dishonesty, a person is typically committed to the thought that it is wrong in that way (for that reason) being the correct thought to have in virtue of that thought meeting a similar but different standard of correctness, again seen as particularly appropriate to the circumstance of judging, according to which a thought about wrongness in that way (for that reason) is appropriate if a person who values integrity (and thereby thinks in terms of wrongness linked to dishonesty) would reasonably judge that it is wrong.

The guiding idea here is that, for example, if we are compassionate then we typically think that a way of thinking about the world that makes no room for a notion like *cruelty* is a failure of that very compassion which, presumably, we value. Why? Because it will typically *seem* like failing to see a moral aspect to actions involving the plight of others undermines the valuing of their plight that compassion essentially involves, by failing to distinguish the significance of their plight from the significance of other things we care about but not in a *compassionate* way. Similarly, if we take ourselves to be people who value integrity (and value that aspect of ourselves) then we will typically see our judgements relating to acts of dishonesty as subject to a standard of correctness to do with the coherence of

³ What is 'reasonable' here? It has to do with being responsive to features of a case which, if the appropriate concepts are in play at all, are supposed to sort cases. For example, if cruelty is in play as a category of moral evaluation, being reasonable in judging things cruel is a matter of knowing what cruelty is supposed to be and whether a particular case has features such as to make it fall under that description.

that valued part of our character and those judgements. In general, our first order moral commitments are, in part, reflections of those virtues of character we take to be part of who we are and value *as* aspects of who we are.

The link between character – specifically our attitudes to our own character – and moral commitment is, on the proposed model, a matter of our moral commitments being linked, *qua* commitments, to those character virtues we take to be partly constitutive of who we are. Thus, a person who values compassion is liable to be committed to moral injunctions against cruelty for the sake of those commitments' relation to that character virtue, and not because of the supposed *truth* of (or of the content of) such commitments, or the good *objective evidence* for them (or for their content). That is to say, a compassionate person for whom their compassion is not taken lightly is liable to be committed to thinking that causing unnecessary suffering is cruel and/or wrong because to think otherwise seems *inter alia* to not properly *be* compassionate.⁴ It might be the case that in fact there is a way to be properly compassionate without being committed to the cruelty or wrongness of some actions; but if that is the case, then it certainly does not follow that it will typically *seem* to be the case to those for whom compassion is important, and the reasons for this are not altogether mysterious – the 'morally thick' concepts such as *cruel* and the deontic concepts such as *morally wrong* afford us ways of thinking and saying things about the world

⁴ It does not seem to follow from a person's taking themselves to *be*, e.g., compassionate that they *value* that character trait in themselves. Be that as it may, in so far as we are concerned here with sincere moral commitment (of the sort typical of ordinary moral deliberation), it is our valuing such perceived aspects of ourselves that is important. On the account I am proposing, utterances or thoughts of moral claims on the basis of character traits one takes oneself to have but does not value are genuine moral utterances or thoughts, but not *sincere* ones: they are not things we *accept* at all, for they are not seen as subject to a standard of *correctness* which they are seen as meeting, but rather just to meet a standard that, in virtue of making reference to a character trait of no perceived *value*, is irrelevant. (Recall that meeting *irrelevant* standards of correctness is *not* a way of being correct in a particular context.)

(including commenting on actions, both possible (as in practical deliberation) and actual (as in evaluation)) in *normative* terms, terms whose relation to action-guiding and evaluation go beyond the resources associated with non-normative descriptive concepts, so it is unsurprising that in taking compassion (for example) *seriously*, as a consideration apt to be influential in practical deliberation and in forming evaluative judgements (what else would we expect of a person 'taking compassion seriously?'), a person is typically liable to think of actions as falling under such concepts, and to resist the rejection of those concepts believing (rightly or wrongly) the rejection of such concepts to threaten the sincerity of their supposed compassion.

This is not a virtue ethics account of morality. It is no part of the account described here that character virtues are in fact what grounds the *truth* of judgements about cruelty or wrongness, or that talk of character is more morally apt than talk of notions such as cruelty or wrongness taken as pertaining to actions or states of affairs (and not in particular to a person's character). *All* that has been claimed is that in characterising what is going on in typical cases of ordinary moral deliberation we ought to acknowledge the importance of a deliberator's *own* concern for particular virtues and for the coherence of their judgements with those virtues. That is just to say that we ought to be sensitive to the fact that when a person accepts that, for example, smacking children is cruel, *that person* is typically committed to a view about the relevance of their own perceived compassion to the acceptance of that claim. Nothing follows merely from *this* about whether the truth or even the objective reasonableness of that claim is tied in any way to that person's – or any other person's – compassion.

At this point it is worth noting a few points about a criticism of virtue ethics accounts of right action that has been nicely articulated by Johnson, 2003. He points out that any account of right action (and, presumably, of

closely related notions) that purports to explain it by reducing it to or identifying it with what a fully virtuous person would do is doomed by the fact that there are things it is right to do which a virtuous person, precisely because they are virtuous, would not do. (One example of this is taking steps to correct a defect of virtue, steps which involve measures distinctive of such correction, which of course a virtuous person would not perform, having no relevant defect to correct.)

The first thing to note is that the account of ordinary moral commitment I have sketched involves no such attempted analysis of right action, nor any such analysis of anything in fact. Nothing in this account is supposed to say anything about what right, or wrong, or cruel etc, actions *are*. The idea is that for all that is said here about ordinary moral commitment, you may take whichever account of what, for example, makes an action right or wrong that you like.

The second thing to note is that it might seem that I have sketched an account of ordinary moral commitment according to which something akin to Johnson's targeted analysis of right is featured as embedded in a person's attitude to some moral claim (for on that account a person *treats* the acceptability of a judgement of an action (or whatever) as depending upon what (who they take to be) a virtuous person would judge), even if not as an account of what is in fact right or cruel or whatever. This is a more reasonable issue to raise, but it is not ultimately problematic. Note first that *even if* it were to be the case that something like the problematic virtue theoretic analysis of right action is featured in my account as embedded in a person's ordinary attitudes to moral claims, that would not be too troublesome: what I am offering is an account of how people often go about ordinary moral deliberation, and it is no part of that account – and there is no obvious reason to expect that it ought to be part of any such account – that such ordinary practice is *rationaly justified* in the way Johnson's point

suggests that it would not be. Many of the ways in which we typically go about ordinary deliberative practices fail to be (fully) so justified in one way or another, and that tells us no more than that our deliberative practices often fall short of meeting the highest standards. This ought not to be surprising. Indeed, the mistake Johnson diagnoses in the virtue theoretic analysis of right action can, in some forms, amount to committing the ‘conditional fallacy’, and as Shope, 1978 points out, many *philosophers*, practiced in the art of careful deliberation if anyone is, have fallen into that particular fallacy. (Shope accuses Roderick Chisholm, Roderick Firth, Gilbert Harman, Keith Lehrer, Norman Malcolm, Michael Slote, W. D. Ross and John Rawls – a not undistinguished group of reasoners!) If the mistake I turned out to be attributing to ordinary moral deliberators were akin to the mistakes made not infrequently in the course of rigorous argument by dedicated practitioners of the arts of reasoning, that would not, I think, embarrass the plausibility of my account of ordinary moral deliberation.⁵

But note, secondly, that it is not the case that something like the problematic virtue theoretic analysis of right action *is* featured in my account as embedded in a person’s ordinary attitudes to moral claims, anyway: it is no part of my account that ordinary moral deliberation typically involves either explicitly or implicitly relying on any principle akin to the virtue theoretic mistake, for I have said nothing about ordinary moral deliberation typically relying upon the thought, on the part of the deliberator, that the *truth* of a claim of right or of wrongness or of cruelty (etc.) depends in any way on that claim being one that a person with the relevant virtue would

⁵ It might seem that no *justificatory* story could successfully involve recommending a *mistake*. But of course there is no reason to assume that the justification offered (elsewhere, in the next chapter, for example) for our practices of moral deliberation and judgement being as pictured here is going to be such as to be incompatible with or undermined by *such* a mistake – a practice does not have to be *without fault* to be apt for recommendation; the issue is just *which* faults are serious enough to undermine justification, and for all that has been said thus far, that remains an open question.

accept.⁶ Indeed, the very point of this account is that a person thinking in the way typical of ordinary moral deliberation is concerned with a moral claim's satisfying or failing to satisfy some standard of correctness which is not truth (or truth-associated) at all – such a person is, rather, concerned with whether that moral claim is the thing to think in some *other* sense. So it cannot be a feature of this account that in accepting a moral claim, such a person is committed to the thought that what makes a claim true is something to do with what a virtuous person would think – they are not in the business of being concerned with the *truth* of that claim at all, but *rather* with the way in which *the accepting of that judgement* is a way of manifesting the relevant perceived virtue (on the part of their own character) which would require (or allow) accepting that claim, whether or not it is true.

We might now see more clearly what the structure of the proposed view is. The proposal places great weight on the attitudes a person has towards their acceptance itself of a moral claim, rather than on the attitudes a person has towards the truth of that claim. Accepting a moral claim, in ordinary moral deliberation, is typically a way of responding to the world in light of considerations to do with the value of being the sort of person who responds in certain sorts of ways. Thus, in accepting that it is wrong to smack children, I would typically be responding to features of smacking children (the discomfort caused, the supposed reasons for doing it, etc.) in a way that seemed to me to be the most sincere and reasonable way of being a compassionate person – I would be guided by considerations to do with, for example, the extent to which I could really be a compassionate person if I were not prepared to think such a thing wrong (or cruel, or whatever).

⁶ I focus on truth here since Johnson (2003: p812) presents the virtue theorist's account of right action as a biconditional, which is most naturally read truth-functionally: the virtue theorist is supposed to be saying (roughly) that it is true that some action is right iff (it is true that) it is the action that a fully virtuous person would characteristically perform.

As noted above, nothing here turns on whether I would be *right* to think, if I did, that being a compassionate person demands thinking that smacking children is wrong (cruel, etc.), nor even on whether I would be right to think that being a compassionate person demands thinking of *anything* in terms of wrongness (cruelty, etc.). That is because the account here is supposed to explain a person's own acceptance of moral claims, and ought therefore to place no restriction upon the possibility of their accepting such claims – such as an 'externalist' requirement that the assumptions involved in their seeing some particular claims as correct in the perceived-as-relevant way be *true*, or that that way really *ought* to be perceived as relevant – which might fail to be met just in virtue of facts outside of the psychology of particular deliberators.

Before looking at some considerations in favour of treating this model as an accurate model of ordinary moral deliberation as it often actually is, it is worth saying something explicit about what precisely is being claimed in respect of what we might call the *mechanism* of moral commitment, and whether my account is overly intellectualist. Clearly, just as beliefs are often involuntary, moral commitments are often unreflective, and it would be entirely implausible to suggest that in incurring a moral commitment the considerations I have adverted to are in every case actually *considered*. But just as with respect to the characterisation of belief I offered as treating some claims as subject to a truth-associated norm and as meeting the standard of correctness associated with that norm, where the 'treating' involved need not be – and had better not be – construed as necessarily *conscious* (so as to allow for unreflective beliefs of various sorts, including tacit beliefs, and/or beliefs we form automatically), the account of moral commitment I am proposing need not – and had better not – depend upon an understanding of what it is to be sensitive to some or other consideration which requires that such sensitivity is necessarily a *conscious* or *reflective* phenomenon.

It might very well be (as much for the account of belief as for the account of moral commitment) that a person's being entitled to claim a reason for their commitment requires that they have explicitly entertained the relation of the claim to which they are committed to the particular standards to which that claim is seen to be subject for its correctness. That is to say, it might very well be that those commitments to the defence of which we are entitled to come are just those commitments about which we have something to *say* in respect of their satisfying some particular correctness condition. But that is not to say that we shan't often find ourselves committed to things about which we have nothing of the sort to say, because our ideas here are not entirely clear to ourselves. I might, for example, have never thought about the fact – though it *is* fact – that I am in the habit of being more sensitive to the relation between a moral claim and my (usually unreflected-upon) self-conception than I am to the relation between a moral claim and my beliefs about what the world is really like. To deny *that* much lack of transparency of our thoughts would seem to be the real face of over-interlectualism.

3. Evidence

The hermeneutic strand of the idea I am developing in this chapter takes the picture of ordinary moral commitment I have drawn and says that that picture is a good representation of the way our actual ordinary moral commitment often is. So besides the mere coherence of that picture, what is required is some reason to think it such a representation.

The considerations I shall adduce in this section are meant to be evidence for the hermeneutic claim. They are not the sort of evidence which entails what they seem to be evidence for; that is to say, they might be misleading evidence. Nonetheless, they amount to some *prima facie* reasons

for adopting the picture I have drawn as a model of actual ordinary moral deliberation, as it often is. And at the very least, they amount to a motivation for the psychological *plausibility* of the model.

One reason to think that the model is psychologically plausible and might very well accurately represent our ordinary moral commitment as is often is, is the way it affords an explanation of the fact that we are often, for example, *ashamed* of or feel *guilty* about our moral commitments.

It might seem as though shame and guilt are characteristically responses to things we have *done*; but if what we do is understood narrowly, so as to include only those (external?) behaviours *which* are chosen or intentionally performed, it ought to be clear that shame and guilt are not in fact so proscribed, for we commonly feel ashamed of having merely *thought* something, or guilty about a choice (the choice *itself*) that we have made, or about the intention itself. I would not like to go so far as to say that our shame or guilt is characteristically directed at our choices, thoughts or intentions themselves even when those choices, thoughts or intentions issue in actions which we also seem to feel ashamed of or guilty about, the shame or guilt putatively directed at the action being, in fact, a response to the choice, thought or intention instead; but I think it is very clear that in many cases we are more interested, in so far as we are concerned with the grounds of shame or guilt, in our reasons for acting, or our sincerity in so acting, than in the action we eventually performed. And to the extent that that is true, the idea that such attitudes are characteristically attitudes towards things we do is right only on an understanding of doing which covers mental actions such as choice and deliberation, and plausibly commitment and judgement as well.

The point I want to go on to make here does not require that we think just about shame and guilt. Though these are the particular attitudes I shall focus on for illustration, the real force of the idea I want to convey concerns

any and all of the 'reactive attitudes' (to employ Strawson's [1962] term), those attitudes which characteristically depend upon seeing their object as the responsibility of a *person*. (This, of course, is essential to Strawson's own point about the reactive attitudes in 'Freedom and Resentment', for it is the fact that attitudes such as resentment require conceiving of their objects as freely performed actions, or free persons, that means we can learn something, Strawson thinks, about the necessity of a metaphysics on which freedom is possible, given that we do in fact have such attitudes. Needless to say, it does not matter at all for our purposes here whether or not Strawson is right about this.⁷) The general point is that the reactive attitudes we adopt towards our own moral commitments, or towards ourselves for being thus committed, and to the moral commitments of others involve linking the commitments in question to our or their character in precisely the way predicted by the account of ordinary moral commitment I have offered.⁸

Suppose that as a young man Jeremy approves of smacking children. He does not take any *pleasure* in the thought of children, even very naughty ones, getting smacked; indeed he thinks that like many things a parent must do in the good raising of children it is regrettable. But it is regrettable

⁷ Darwall, 2006 has made use of this feature of the reactive attitudes in developing his 'second-personal' account of normativity, and argued that Fichte and Dewey amongst others ought to be read as focussing on the interpersonal recognition of person-hood that the second-personal standpoint and the reactive attitudes involve. Again, the correctness of this is, for our purposes, not at issue. But is it clear that the reactive attitudes are morally interesting for more than just their being attitudes we adopt on the *basis* of moral commitments – they seem (or at least have seemed) to be instructive with respect to the *nature* of moral commitments as well.

⁸ It is not uncontroversial, of course, that shame in particular is an attitude directed in a negative way towards the self (see, e.g., Velleman, 2001 or Williams, 1993 (esp. Appendix 2) for accounts of shame which are not committed to this, and examples supposed to show that no such commitment could be right). But I am inclined to think that many if not all of the case-studies of 'shame' which are supposed to cast doubt on this (such as the 'shame' felt by a life model upon noticing the glint of lust in the artist's eye, or the 'shame' attributed to Adam and Eve upon learning of their nakedness) are in fact better described as cases of *embarrassment*; unless, of course, the sort of negative self-assessment I am suggesting is necessary for shame is stipulated to be felt in addition to the obvious embarrassment such situations involve.

because it is unfortunate to have to hurt a child, even a little bit, which is a slightly upsetting thing to contemplate, not because it is *morally* regrettable. In fact, the young Jeremy might be of the opinion that it is *morally* regrettable to *not* smack children on the very rare occasions when so treating them would be genuinely beneficial to their moral and personal development. At this stage of his life, though, Jeremy does not have children of his own, and he would never dream of smacking someone else's children (it is, even when justified, always a *parent's* prerogative *cum* duty). So, whilst he is fully committed to the thought that smacking children is permissible and, in particular circumstances, perhaps even required, he remains committed to this in thought only, never having occasion (nor regretting his lack of occasion) to smack a child himself. And being of the opinion that parenting is the sort of thing it is best not to advise others on (particularly since he has no experience), he never encourages others to smack their children (or anyone else's), either.

Nonetheless, middle age finds Jeremy with a rather different attitude towards smacking. He still has no children, but he now thinks that smacking children is wrong (perhaps because he now thinks that it is cruel). And he is ashamed of having ever thought that it is morally permissible and, even worse, sometimes morally required to smack one's child. He is obviously not ashamed of having smacked anyone, for he never did, nor for having been a material accessory to any smacking, for he was never was (since he kept his opinion to himself). Nonetheless, it is enough that he even *thought* that smacking children is right to make him ashamed of his former self.

Jeremy might have changed his mind due to having learned some new facts about how distressing being smacked is for a child (or how distressing smacking turns out to be for parents), or how effective in achieving behaviour modifications smacking really is. If that were the case, Jeremy would be liable to see his earlier commitment to the rightness of

smacking as akin to (or even just a species of) an ordinary mistake of judgement in which he just didn't know all – or enough of – the facts, and would be liable to *regret* what he has come to think of as a mistake – as he regrets not having known all the other things he has learned with increasing maturity – but not, just in virtue of that former lack of knowledge, to feel *shame*. He *might*, of course, be liable to feel shame if he were of the opinion that his former lack of knowledge was his fault, an abdication of epistemic responsibility, for example, though it is rarely *shame* that attends even the acknowledgment of culpable epistemic failings. (Surely we more usually regret our failing and resolve to do better, rather than evaluating ourselves in the distinctively negative way characteristic of shame.) But *that* sort of response to a former self's mistake, even if it were to amount to feeling ashamed, would be a feeling of shame directed not at his former moral commitment but at his performance at empirical data collection, or his haste in forming opinions on insufficient evidence.

As it happens, though, it is no such new information which is responsible for Jeremy's change of view, and his feeling of shame is directed at his having been committed to that moral view, not at his failing to properly collect evidence or at his haste in forming opinions. In fact, part of what makes him so ashamed of his former commitment is precisely that he knows very well that he had just as much information then as he does now, and cannot therefore absolve himself on grounds of being misled. What, then, accounts for his shame at having been thus committed? Appealing to the resources available from the account of ordinary moral commitment just described, we can say that Jeremy has just come to see the smacking issue in a different way, though his stock of information about the mechanics and effects of smacking is just the same as when he was young. It now seems to him that what really matters in life – at the very least, what matter to *him* in *his* life – is to be compassionate, to be really properly concerned with the

plight of others, and that *that* is more important than the demands of some abstract theory of human good according to which the odd well judged smack is a permissible or even required means to the end of helping a child to grow into a responsible, temperate adult. He has *not* come to think of that idea of human good as *wrong* in middle age; he has just come to see it as *irrelevant* with respect to deciding what he ought to say or think about the issue of how one person (a parent) ought to behave towards another (their child). He has come to think, rightly or wrongly, that his interest in compassion requires him to condemn anything that looks much like violence to another human being, at least when such violence (albeit relatively low level) serves no broader humanitarian goal. He knows, as well as he ever did, that strictly speaking there are things of value in life beyond freedom from even low level pseudo-violence, and that in order to maximise those values it might *really* be okay to accept a pay-off such as is involved (he might still think) in smacking a child to make them a better person. But he does not want to be the sort of person who calmly and detachedly approves of smacking for *that* reason, and he is ashamed of having *been* that person in his youth, for such a person, he thinks, rightly or wrongly, might have the virtue of being right but lacks the virtue of compassion, and compassion is, for him, with respect to these sorts of issues, the really important virtue. (We might put this by saying that the middle aged Jeremy would rather be a kind fool than a clever knave; a similar, though in reverse, point is encoded, I think, in the common thought that a young man of the right is heartless whilst an old man of the left is stupid – the point of that thought, I take it, is to do with the salience of the different vices at different times of life.)

Thus Jeremy's middle-aged shame is bound up with the way in which what he accepted as a young man was accepted on the *wrong grounds*, or on *inappropriate* grounds, and the specific sort of mistaking the grounds on which to accept claims about smacking involved is to do with having lacked

sufficient interest in compassion to lead his former self to seek a moral commitment in line with the demands of *that* rather than the demands of truth.⁹

It might seem that the involvement of a former attitude that at least *looks* doxastic, on my account of the doxastic/nodoxastic distinction, might be problematic here. But other examples will do just as well for the purpose in hand. Suppose Boris used to think that driving when drunk is permissible, not because he thought it really *is* (he thought even then that the right moral theory probably condemned it), but because he didn't *care* about what he really ought to do, accepting moral claims instead on the basis of whether or not they seemed to be the thing to think for a person like him who had a healthy sense of fun and of youthful disregard for the rules. Now, older and wiser, he is ashamed of having thought that driving when drunk is permissible, even though he never did any driving (drunk or sober) anyway. He doesn't think he made an epistemic mistake – he knew then just as he knows now how dangerous it is, and how no moral theory that is strictly true could condone it – but he now recoils at the thought of having been a

⁹ Independently of the plausibility of the model of ordinary moral commitment I have suggested, this is not, I think, a far-fetched thing to say about our common attitudes to our former moral commitments. The basic truth in it is manifested also, I think, in truisms about the 'idealism of youth'. Such truisms are not plausibly understood, if they are to encode any truth at all, as saying that the principles to which the young are often attached are *false*: it is not the *stupidity* of youth that is being adverted to here. Rather, the interesting idea conveyed by such truisms is that *whether or not* the principles to which the young are often attached are right, their attachment to those principles is not of the right sort: it is commitment driven entirely by supposed truth and understanding, insensitive to the more germane grounds of commitment appropriate to the business of real life moral thought. This idea is commonly conveyed by talk of age bringing a more 'realistic' attitude to morals, but it is striking that by this is rarely meant a more *literally accurate* attitude – witness arguments between idealistic youths and their wiser elders, in which the elders tend, invariably, to terminate their arguments not with attempted refutations of the youth's preferred principles, but with withering deprecations of the *point* of those principles in the real world.*

* Nothing here, or anywhere in this thesis, turns on the fact that I have invoked *principles* – such exchanges might, I suppose, be taken by some as grist to the mill of a hermeneutic form of *particularism*; but I think it is reasonably clear that the truth in the truisms about the idealism of youth depends upon the type of moral certainty involved in that idealism, and not upon the nature of what that idealism is certainty of.

person for whom the acceptance of such a claim turned, not on acceptance of that claim being appropriate for the sake of taking a sober and admirably cautious attitude to the world (as he now does), but on whether that claim was required by that immature attitude to the world he has thankfully grown out of. Here, neither Boris's former nor his latter commitment is based on his view of the truth or otherwise of the literally construed moral claim that it is permissible to drive when drunk: in each case, he is concerned not with that but with what, *qua* the sort of person he is and values being, he ought to think, and in neither case does *that* involve being the sort of person who only thinks what is true, or probably true, or well supported by evidence. The younger Boris would have seen truth as boringly irrelevant – he wants to be the sort of person who thinks that driving fast cars in an inebriated state is morally permissible, because he wants to be the sort of person for whom thinking that sort of thing is an appropriate way of seeing the world in permissive, rebellious terms. And the older Boris would be just as suspicious of the relevance of truth, for he would be much more concerned with being the sort of person who treats such irresponsibility as wrong, because that is an appropriate way of manifesting the sort of cautious, sensible attitude he (now) thinks fitting. Just as in the case of Jeremy's change of mind, though, Boris is liable to feel ashamed, and what he is liable to feel ashamed *of* is his having been *the sort of person* for whom acceptance of moral claims depended on youthful foolishness rather than mature respectability.

Other explanations of Jeremy's shame and of Boris's shame are available. But one virtue of my suggestion is that it does not require us to insist that Jeremy and Boris must have learned something new. This is a virtue since, *ex hypothesi*, they have learned no new empirical facts. They might, as some accounts would have it, have learned the *significance* of some empirical facts they already knew, that is true. But that, construed as a

change of *belief*, would seem to require that, if they are to be credited with at least as much consistency as they had before, they must now have a new *theory* of moral significance which, *ex hypothesi* they do *not* have – they still think that smacking or prescinding from drunk driving really are required by the moral theories they believe to be correct (they just refuse to see that as the appropriate standard of correctness in ordinary moral deliberative contexts). And of course the fact – if it is a fact – that my account of *Jeremy's* or of *Boris's* shame at their former commitments is plausible and lends support to the general account of ordinary moral commitment that I am offering does not show once and for all that that general account is implicated in plausible explanations of *all* such cases. But the point I want to make with these examples is that, firstly, it is not at all uncommon to adopt attitudes such as shame towards our moral commitments themselves and, secondly, that in cases where we do, our attitude is typically directed not at either our epistemic performance or our merely having been committed to the wrong thing, but rather at being the *sort of person* who incurs that sort of commitment in that particular sort of way. My account of ordinary moral commitment explains that phenomenon in the most *economical* way, since on it a person's attitude to the sort of person they are (or would like to be) is involved in their acceptance of moral claims from the very first – attitudes to the role of one's *character* in accepting particular moral claims (such as the attitudes involved in the reactive attitudes) are not, on my account, required to latch onto anything (that very role of character) which isn't already of the very essence of moral commitment. Put another way, my account does not require us to say that a person's seeing their former moral commitments as having involved character in an undesirable way is *revisionary* or mistaken, for ordinary moral commitment typically *does* involve reflecting our character in our commitments.

Nonetheless, attitudes to character are obviously involved in assessments of our commitments of all sorts, including our ordinary beliefs. In such cases, it is our epistemic performance, and our characteristic traits which explain that performance, which might be the objects of our negative assessment. But that seems to mean that there is no straightforward argument from the shame-aptness of *moral* commitment to the essentially nondoxastic account of that commitment I have given. It is worth noting that there is in fact a way to go in response to such an observation so as to *accommodate* the data upon which it relies: we might say that the straightforward argument from shame-aptness to nondoxasticism is sound, and that that just shows that our ordinary commitments of *all* sorts are not so doxastic as we first thought. That response would not need to go so far as to embrace *universal* nondoxasticism, at least if the reach of its conclusions were limited to the *ordinary*. And such general nondoxasticism may be true – at least I don't wish to argue here that it is not. But it would be a blow to the plausibility of the nondoxastic account of ordinary moral commitment if that account were to entail such general nondoxasticism. So something ought to be said about why shame-aptness is apt to play a role in an argument for nondoxasticism about ordinary *moral* commitment, but not about our ordinary commitments of other sorts.

The way to drive the required wedge is, I think, to notice that it is not mere shame-aptness which is pulling the load here. As noted in discussing Jeremy's case, we might feel ashamed at being, for example, too lazy to have bothered to find out enough to base a reasonable opinion on, or at being the sort of person for whom taking a sufficient interest is just too boring. But our shame at having thought something the thinking of which is attributable to those character flaws is not characteristically the shame of a vicious self-conception. Moral shame, when directed at our moral commitments rather than our behaviour, is characteristically shame at having been the sort of

person for whom the thing we thought *is* the thing to think. We feel bad, in a way, because we made no mistake about *that*. But for someone who looks back on their former epistemic irresponsibility, it is not their being someone for whom there is no mistake in drawing the conclusion they drew that makes them ashamed – they are ashamed for having made a mistake of just the sort which, had they reflected at the time on their performance, they would have diagnosed *as* a mistake. That is to say that their lack of epistemic virtue (perhaps their laziness, or lack of sensitivity to the relevance of associated issues or evidence...) would, at any time, have struck them, where they made aware of it, as a fault, and would certainly not have seemed to them to be a *reason* for – though it may be the *cause* of – their accepting a particular claim. Thus the shame we feel when coming to see our epistemic performance as vicious in this way is bound up with our having manifested a vice which led to our being mistaken, but not with that vice having been a character trait which we – shamefully – *embraced*. However benighted our belief forming mechanisms are, to the extent that we are genuinely forming *beliefs* we are taking our commitments to be fully correct (at least) only if true, and any epistemic vice which threatens our being committed only (at least) to truths cannot *thus* far be first-personally endorsed. So in feeling ashamed of having manifested those epistemic vices, we are not ashamed of having endorsed those vices as grounds of belief formation; we are merely ashamed of having manifested those epistemic vices. It is the shame of having embraced or endorsed – perhaps implicitly – being a particular sort of person (or of *not* having embraced being a particular sort of person) that is distinctive of moral shame, when it is our commitments themselves that we are ashamed of.

So the point is that moral shame at having been committed to particular moral *thoughts* is character directed, as is all shame; but unlike the shame we feel at having made, for example, epistemic mistakes, it is directed

at not only the vices we in fact now seem to have manifested and the causal role they are now seen to have played in our being thus committed, but also at our having *endorsed* those vices (implicitly or explicitly). This is not to undermine the point I was concerned to make about moral shame being a matter of deprecating one's own character, for it is surely a matter of one's character whether or not one was disposed to endorse some character trait. Indeed, it is just *such* elements of character that the account of ordinary moral commitment I have proposed predicts will be the object of moral shame, for it is the endorsing of perceived virtues which, on that account, demands or allows the acceptance of particular moral claims.

So much for shame and the reactive attitudes. I want to now say something about the way in which the account I have given of ordinary moral commitment can shed light upon another phenomenon of interest to theorists of ordinary moral psychology, namely the connection between moral commitment and motivation. Any account of ordinary moral commitment ought to explain the extent (and limits) of that connection, the interest of which arises from the seeming fact that in accepting a moral claim in the way distinctive of ordinary moral commitment we are typically thereby in possession of a good reason to act in particular ways (should relevant circumstances arise), which is not generally the case with respect to our acceptance of other sorts of claims. One attractive way to approach this interesting phenomenon is to see it as having to do with the distinctively *normative* character of *what* is accepted in accepting moral claims (the normativity, that is, of moral claims themselves). This in turn might or might not be understood in terms of the normative character of what those claims are *about*, in terms, that is, of normative facts. But the prospects for understanding the normative *implications* for practical reasoning of accepting moral claims by means of an account which bottoms out in a story about the normative facts *described by* moral claims are not obviously promising: there

are, plausibly, various sorts of normative facts described by claims of various sorts which do *not* have the interesting connection to motivation that moral claims seem to display; and even if that were not the case, such an account would seem to be the wrong *sort* of explanation, missing its mark of explaining what the *role* of our normative commitments are regardless of their *content*. The prospects for an alternative, *dispositional* account (in the manner of, for example, Wedgwood, 2007), which settles for claiming it to be a psychological fact about us that our acceptance of normative claims tends to amount to, or to be typically correlated with, a motivational state, seems to depend entirely upon how the disposition in question is spelled out.¹⁰ It is just such a spelling out which, I want to suggest, my account of ordinary moral commitment can provide.

It ought to be no surprise by now that my suggestion is that the involvement of seeing oneself as a particular type of person, and seeing the perceived demands of being that type of person as the relevant standard of correctness for accepting moral claims, is central to the account I want to give of *why* we are disposed (if we are) to treat moral claims that we accept as reasons, if we are rational.

How precisely that works is not, in fact, all that important here. All that is required in order for the motivational oomph of moral commitment to

¹⁰ Is it a *psychological* fact, if we are supposed to intend to ϕ if we accept that we ought to ϕ , *if we are rational* (as Wedgwood (2007: ch1) has it), or is that a fact about rationality instead? Cf. e.g. Foot, 2001: ch4, here, for the idea that there is not a clear distinction to be drawn between facts about rationality and facts about deliberation. Wedgwood, in defending his preferred version of internalism (Wedgwood, 2007: p27), suggests that the disposition in question might be attributed on Davidsonian methodological grounds (for reasons bound up with charitable interpretation and the attribution of mental states), by which he means that attributions of rationality are already bound up with attributions of such dispositions. (Wedgwood does, however, return to the issue with his own theory in Wedgwood, 2007: ch7.) But of course that doesn't get us any closer to understanding what having such a disposition *amounts to* in the sense of what explains the psychological facts – all the Davidsonian grounds get us, if we are sympathetic to those grounds, is that we must attribute such dispositions, *whatever may explain them* (though a thoroughgoing Davidsonian would be suspicious of what further there could be to say about that anyway).

be grist to the mill of my account is that what does the explanatory work is the role in such commitment of *self conception* (which, of course, need not be taken to require explicitly entertaining any ideas of one's self, and which is not to be understood as merely descriptive but also as involving attitudes to the self one *wants* to be). The idea is that my account of the character of ordinary moral commitment is such as to make it obviously irrational to *not* count the acceptance of a moral claim as a *prima facie* reason for acting on it, or for intending to act on it. Of course there are various ways in which it is rational to act on a particular claim that we accept, and it will not always be (most) rational to act in the way that, for example, a moral claim we accept *says* we should act (for we might, on pretty much *any* account, have good countervailing reasons to act otherwise, and might be aware of those reasons). What is irrational is, minimally, to not see our acceptance of a moral claim as demanding a role in our practical deliberation. That is to say that the irrationality involved in accepting a moral claim without any disposition to intend to act on it is the irrationality of thus accepting claims which demand to play a role in our practical deliberation, *somehow*. But why should we think that moral claims are such as to demand just that? That is the question the answer to which will amount to something approaching an adequate spelling out of the idea that rationality demands a dispositional connection between our acceptance of moral claims and our motives. And the answer to that question, on my view, is that practical deliberation, in the sense relevant here, is not just (or not *even*) deliberation about '*what should be done*', but deliberation about '*what I should do*', and that in deliberating about what *I* should do it is rational for me to give weight to commitments I have *in virtue of* me being the person I am and want to be: failing to give weight to *those* commitments is failing to treat them as what they are – namely (*inter alia*) expressions of the sort of character I value having. For if I fail to give weight to such commitments in practical deliberation, I effectively overlook a

valued aspect of my (perceived) character, and I therefore embark on a decision about what I should do that gives less weight than might have been given to considerations which would tie my practical conclusions to my understanding of who I am. And that would be irrational in so far as the essentially first-personal character of practical deliberation implies a conception of rationality for such deliberation which involves consistency between our understanding of ourselves and our decisions about what to do.

The reason *beliefs* are not typically linked to motives in the interesting way is that my believing something is not characteristically such an expression of my self-understanding. I *might* have come to believe something because of the sort of person I am, and might know that, but it is not of the very nature of beliefs that they are such as to require us, for the sake of *practical* rationality, to be disposed to tie our beliefs to our decisions about what to do in the relevant sense, for whilst it is a sin of practical rationality to ignore relevant beliefs, it is not a sin of *practical* rationality to fail to intend to act on them, for I might have good reason to place no weight on some relevant belief (would that just render a putatively relevant belief *irrelevant*? plausibly not...) having considered it. This is not so with our ordinary acceptance of moral claims, on the account I have offered, for we can never have a good reason to see relevant claims which express something about our character which we value as having *no* weight in the context of a deliberation for which any reasonable conception of rationality has to do with the coherence of our conclusions with our self-understanding.

These considerations – addressing my account's relation to the phenomena of the moral reactive attitudes and the explanation of the motivational relevance of moral commitment – have been largely matters of theoretical adequacy: given the need to account for some acknowledged phenomena, my account can provide simple and intuitive resources with which to meet that need. The real plausibility of the hermeneutic strand of

my account probably rests on the extent to which it is seen to capture something of the essence of ordinary moral commitment in an intuitive way, though. It is, of course, for the reader to decide that for yourself. But in thinking about that, it is important to be aware of the need for sensitivity. We ought not to assume, for example, that when people commit to moral claims at odds with what we have heard them commit to the literal falsity of in other contexts that they are either being straightforwardly inconsistent or insincere, or that they have just changed their mind. We ought to be alert to the wider possibilities for explaining the commitments people tend to manifest, and to look for *charitable* accounts of that. And that demand for charity ought not to be read in too narrow a way: we ought not to think that the only way to be charitable in interpreting commitments is to find a way of making a person's commitments *logically* coherent – sometimes a charitable interpretation of their commitments, taken as a whole, will require treating them as jointly reasonable in some *other* way.

4. Excursus: A Literary Illustration

The passage from Dostoevsky's *Crime and Punishment* which forms the epigraph of this chapter is put in the mouth of Andréy Semyónovitch Lebeziátnikov, the friend of Pyótr Petróvitch Lúzhin, whom Lebeziátnikov wrongly believes has just slipped a substantial sum of money surreptitiously into the pocket of a destitute young woman out of benevolence. In fact, Lúzhin's action was not benevolent at all, his purpose being to eventually suggest, in front of a crowd, that the unfortunate woman had stolen the money and thereby had it about her person. And Lebeziátnikov (whose name is derived from the verb *lebeziat*, meaning to fawn or cringe) is portrayed as something of a fool, 'one of the numerous and varied legion of dullards, of half-animate abortions, conceited, half-educated coxcombs, who

attach themselves to the idea most in fashion only to vulgarise it and who caricature every cause they serve, however sincerely' (Dostoevsky [1866]: p308), an uncritical ideologue committed to the socialist doctrines of the left-wing political agitators of Petersburg's radical underground. But neither the unsympathetic characterisation of Lebeziátnikov himself nor his mistaking what he saw undermine the illustrative point of invoking this passage, which is to highlight the way in which it can seem very natural to us for a person to judge an action favourably from a moral point of view, *even though they are committed to a theoretical stance according to which that action is not morally right*. Lebeziátnikov is, foolishly or otherwise, committed to the idea that private benevolence is deleterious to the elimination of social ills and thus, as a good radical socialist, to either the idea that there is no such thing as the morally virtuous (if he is suspicious of the very idea of moral evaluation, preferring to replace it with 'scientific' notions of efficiency and the like) or the idea that, as socially divisive, such benevolence is morally *bad*. Nonetheless, he is not evidently prepared to let those commitments, which we might think of as commitments of his considered moral theory (to the extent that he is represented as considering anything very deeply), stand in the way of praising what he takes to be Lúzhin's benevolence in what we might think of as his *ordinary* moral mood, the thought about actions and their virtues or deontic status that he employs when away from his books or discussions of moral and political doctrine and engaged in the everyday cut and thrust of praising and blaming.

To be sure, he is careful to remind Lúzhin of his (Lebeziátnikov's) 'official' stance on benevolence, even as he praises him (Lúzhin) in direct contravention of the demands of that stance. But if Lebeziátnikov looks ridiculous in doing this, it is *not* because it is ridiculous that he should praise Lúzhin for what he takes to be an act of benevolence, given his other commitments; it is rather that he should invoke those other commitments at

all when praising Lúzhin, for what seems out of place is not his positive evaluation of Lúzhin's action (or what he takes to be his action) – for what could be more reasonable than thinking well of a man for an act of benevolence? – but rather his eagerness to use the occasion to remind Lúzhin of his (Lebeziátnikov's) considered views on the *true* moral nature of benevolence.

Note that it would not be right to represent this as a case in which what is ridiculous about Lebeziátnikov's speech is his *posturing* at a stance on which benevolence is *in fact* not morally admirable, a mere posturing which is revealed as such by his *actually* praising the perceived benevolence of the action is question – Lebeziátnikov is, throughout, portrayed as deeply committed (foolishly, perhaps, but committed nonetheless) to his theoretical cause, as evidenced by his genuine irritation and anger at Lúzhin's refusal to take those commitments seriously. And whilst Lebeziátnikov is portrayed as a dullard, it would be deeply uncharitable to see his praise and his explicit claim that such praise is in tension with his theoretical commitments as borne of genuine confusion about whether he really is committed to contradictory things. It seems far more reasonable to acknowledge Lebeziátnikov's sense that his various sorts of commitments – the 'thoughtful' and the ordinary – really are in some sort of tension.

Literary examples such as this are instructive in this sort of philosophical context because (and only to the extent that) they strike us as realistic. If Lebeziátnikov is unrecognisable in respect of the feature I have been drawing attention to then there is no evidence for my suggestion to be found in examining his behaviour here. But if we think that the way Lebeziátnikov behaves here *is* recognisable (modulo, of course, his *particular* theoretical commitments, and the particular situation in which we find him with respect to *Lúzhin's* actions) then we have reason to be sympathetic to the suggestion I want to make about the character of ordinary moral

commitment. For Lebeziátnikov is manifesting precisely the separation of ordinary moral thought and theoretical commitment that I have said is a feature of moral discourse, and his ordinary moral evaluation of Lúzhin's action is associated with his positive evaluation of the character trait of benevolence, *despite his commitment to benevolence not literally being a moral good*, in just the way my suggested account would predict: Lebeziátnikov values benevolence not because he believes it to be a moral good (he explicitly does not), but because he feels well disposed towards it nonetheless (note his language here: 'I saw your action with *pleasure* – yes, yes, I *liked* it' (my emphasis)). He congratulates Lúzhin for doing a good thing, though he does not believe that what he has done *really is* a good thing. His ordinary moral commitment to it being a good thing is not undermined, though – the warmth with which he addresses Lúzhin is testament to that. Clearly there is something interesting going on here, more interesting than mere inconsistency. What is being manifested is a commitment to the value of benevolence, and the recognition that what Lúzhin has done (or what he is supposed to have done) accords with the demands of benevolence. It is not that what has been done is *believed* to be a morally good thing – from Lebeziátnikov's point of view, what has been done, *qua* act of benevolence, is *not* a morally good thing, properly speaking – but rather that there is some other sort of commitment to that being a good thing (a morally good thing, an 'honourable' thing), a commitment which is based not on accordance with what is taken to be literally true but rather what is taken to accord with the demands of a virtue of character which, though not literally *morally* good, *just is* valued in some way. And of course it is only possible to make sense of all this if we understand that Lebeziátnikov's *own* endorsement of benevolence (as a character virtue, if not as a sort of action) is at the root of his praise of Lúzhin's perceived benevolence.

This is just the sort of charitable understanding of a person's manifest commitments which I suggested, above, is necessary and which, when carried out with due sensitivity, seems likely to support my claim that our ordinary moral commitments often are, or at least quite easily *could be*, as I have described in this chapter. In elaborating on this *literary* example in the interests of adducing evidence for my claim about actual moral deliberation and commitment, I am not supposing that there is anything distinctive to be gleaned from the fact that this *is* a literary example. I am, in fact, on the side of Raphael, 1983 in seeing the moral value of literature as bound up with its capacity to bring certain facts to salience or to highlight interesting dimensions of cases etc., against Crary, 2007 (esp. ch4) and others who suppose that there is something (philosophically) ineffable in the moral lessons available from literature. And I think that, similarly, in metaethics we can usefully appeal to literature in the same way as we can appeal to history and the other sources of knowledge of human nature in order to widen our evidence base. Which is not, of course, to allow literature any *special* place in the theory of moral psychology, nor to allow that there is any role for it to play in such theories which is not tied to and constrained by the extent to which we recognise it as realistic.

5. Conclusion

Moral commitment is often nondoxastic. On the model I have described and motivated, which for want of a better name I shall call the 'Character-Centred' model, our nondoxastic moral commitments consist in the acceptance of moral claims or principles on the basis of their being seen as required or permitted by the demands of being the sort of person we want to be, or of manifesting what we see as virtues. Nothing in accepting a moral claim or principle on *that* sort of basis involves seeing the claims or

principles accepted as true, or as subject to the moral facts for their correctness.

In the next chapter, I shall argue that there is reason to think of it as a *good* thing if our moral commitments often are as the Character-Centred model represents them to be.

The Good in the Nondoxastic

There is another Humour, which may be observ'd in some Pretenders to Wisdom, and which, if not so pernicious as the idle petulant Humour above-mention'd, must, however, have a very bad Effect on those, who indulge it. I mean that grave philosophic Endeavour after Perfection, which, under Pretext of reforming Prejudices and Errors, strikes at all the most endearing Sentiments of the Heart, and all the most useful Byasses and Instincts, which can govern a human Creature.

David Hume, 'Of Moral Prejudices'¹

1. Introduction

What would become of us if our acceptance of moral judgements and commitments were typically some attitude other than belief, even though the contents of those judgements were apt to be believed – if, say, our moral judgments and commitments were as described in the previous chapter? If one is a *hermeneutic* fictionalist, then the question is: what *shall*, or what *has*, become of us, since our acceptance of moral judgements *is* typically some attitude other than belief, even though the contents of those judgements are apt to be believed? For the *revolutionary* fictionalist, the question is: why *should* we revise our existing typical practice(s) of moral judgement so as to have moral judgement or commitment typically amount to some attitude other than belief, even though the contents of those judgements would still be apt to be believed? In both cases, the lurking suspicion is that as long as the contents of our moral judgments are *apt* to be believed, they *ought* to be believed if accepted at all, on pain of... well, on pain of what?

¹ Hume, [1752]: p539. The 'Humour above-mention'd' is, seemingly, the tendency to elevate personal enjoyments and passing fancies above such stabilising virtues as respect for and benevolence towards family and country (pp538-9).

One worry might be that our typical moral judgements are bound to lack motivational oomph if they fall short of beliefs, and that would be a bad thing. After all, we are not likely to be strongly disposed to act on hard moral choices if those choices terminate not in belief but in some other sort of commitment, akin to our acceptance of fictional claims, or so it might seem at least. I shall argue, however, that the reverse is true – that if our typical moral judgements *are* beliefs then they are bound to lack at least the *security* of oomph that is desirable.

My argument will be based on a feature of what I'll call *subjective epistemic warrant*, and will turn on common features of moral beliefs, and their implications for action-guidingness. But aside from the positive argument that there is reason to be suspicious of the adequacy of moral *beliefs* to secure action in cases of hard moral choice, and that there is reason to think the sort of nondoxastic commitment I have in mind to characterize moral judgement would do better, it will be worth noting that there is nothing inherently flimsy in the notion of nondoxastic commitment, generally.² In the next section, I shall give some examples of seemingly very robust nondoxastic commitments that I assume most of us have.

In Section 3 I say something more about the sort of 'evaluative fictionalism' I defend in this paper, stronger than hermeneutic fictionalism but weaker than revolutionary fictionalism, and about a *prima facie* reason one might have for suspecting that any sort of fictionalism about matters

² It might be best for the purposes of this chapter to focus on moral *judgement* rather than commitment. See Chapter Four, Section 1, for a brief account of my views about the relation of moral judgment to moral commitment and moral deliberation. As I have said, I think that, in the absence of compelling reason not to at least, we should see a nondoxastic theory of moral *commitment* (such as the one developed in Chapter Four) as implying a related nondoxastic theory of moral judgment and deliberation. But perhaps focussing on judgment would have the virtue of respecting the thought that hard moral *choices* (with which we shall be concerned) are, *qua* choices, a species of judgment. But it seem perfectly reasonable to think that the action guiding status of our *commitments* is just as amenable to the sort of considerations and argument I shall be presenting as our moral judgments are, so I shall not restrict myself to talk of judgements in what follows.

pertaining to practical reason (such as moral judgement) might be hopeless. In Section 4 I give my Subjective Warrant Argument for the superiority of nondoxastic moral judgement over moral belief in respect of general rational action-guiding potential. And in Section 5 I suggest how *realistic* fictionalism, if accepted, would undermine at least one sort of worry about that argument

2. Must Nondoxastic Commitments be Flimsy?

There is no reason to suspect nondoxastic commitments of being at all 'flimsy' in the sense of lacking the robustness of beliefs. For one thing, the idea that they are bound to be flimsy – lightly adopted, non-action-guiding, easily given up or revised – is apt to seem rather theory laden: if we start with the idea that some of our most important commitments, such as our moral commitments, *are* robust, and that that is partly what makes them interesting, then we have pre-theoretical reason to keep an open mind about the possible bearers of robustness at least until we have satisfied ourselves on *other* grounds that in fact no nondoxastic commitment is robust. Put another way, if it is an open question whether or not our moral commitments, for example, are doxastic, but we know that they are interestingly robust in any case, we had better not say that we already know that no nondoxastic commitment could be robust.

For another thing, it is easy to think of nondoxastic commitments which are obviously quite robust. It is not at all obvious – in fact it seems false – that our 'playing along' (if that is what is going on) with Sherlock Holmes being a detective is something we are liable to just opt out of without good reason. Nobody, least of all the fictionalist, is committed to saying that there are no *reasons* for nondoxastic attitudes, and the demand that good reasons be found for changing those attitudes is no less plausible than the demand that, for example, belief bear some normative relation,

which a person might be strongly inclined to respect, to evidence. The idea that a person would be more easily persuaded to sincerely assert that Sherlock Holmes was *not* a detective than to change their mind about some robust belief is really quite puzzling – surely, a person would rather say that they don't know what to say about Sherlock Holmes at all any more than say, with Sherlock Holmes in mind, that he was not a detective, and it is hard to see what could convince them that they did not know what to say about Sherlock Holmes in respect of his being or not being a detective: if Sherlock Holmes is the object of *any* attitude (in a literary context), then surely he is a detective.

Clearly not *everything* we thus accept is so robust. But it needn't be. All that is required to resist the general form of the challenge that nondoxastic attitudes are bound to be flimsy is that *some* nondoxastic attitudes are sufficiently robust to count as genuine commitments in the sense of being given up (ideally) not at all lightly, if that is the relevant mark of robustness. Nor does adverting to this sort of example of the robustness of nondoxastic attitudes mean the renunciation of the point that the fictionalist's relevant nondoxastic attitude is not necessarily anything to do with fiction, for all the example was supposed to do was to falsify the thought that nondoxastic attitudes are bound to be flimsy, from the falsification of which it follows that there is no *general* problem *in re* robustness for a nondoxastic account of, for example, morality.

3. Fictionalisms: Hermeneutic, Revolutionary, and Evaluative

I have tried so far, in laying out a brief taxonomy of fictionalist options and my preferred account of ordinary moral commitment, to remain neutral between 'hermeneutic' and 'revolutionary' forms of fictionalism, attempting to use talk of pictures, models and the like to describe a fictionalist *position*

without saying what it is a position concerning (what the pictures or models are *of*). ‘Hermeneutic’ fictionalism is a view about what our acceptance of claims of the target discourse (typically) *is* like. ‘Revolutionary’ fictionalists deny hermeneutic fictionalism but propose a change in our (typical) practice to bring it into line with what the hermeneutic fictionalist thinks that practice is like already.

It is usual for discussions of fictionalism to recognise the distinction between hermeneutic and revolutionary sorts of fictionalism, but really a finer grained taxonomy is required. The revolutionary fictionalist *recommends* that our acceptance of their target claims be (typically) nondoxastic, and that means (if the revolutionary fictionalist is reasonable) that they think there is some reason to approve of our acceptance being (typically) nondoxastic.³ But *that* thought (that there is some reason to approve of our acceptance being (typically) nondoxastic) is a thought that might be had by any of three types of theorist. Of course it might be thought by a revolutionary fictionalist – indeed, it is plausible that *principled* revolutionary fictionalism entails such a thought. It might also be thought, though, by a hermeneutic fictionalist who thinks that the way our (typical) acceptance of their target claims *is* is actually a *good* way for it to be. But the thought that there is some reason to approve of our acceptance being (typically) nondoxastic is obviously not entailed by the thought that our acceptance *is* (typically) nondoxastic. So the thought about the *merits* of nondoxastic acceptance which is potentially shared by hermeneutic and revolutionary fictionalists is apt for attention itself. Call someone who defends the *merits* of nondoxastic acceptance an *evaluative* fictionalist. (Here, ‘evaluative fictionalism’ is not fictionalism about

³ Throughout, it is to go without saying that *any* sort of fictionalist is not just concerned with the (typical) mode of our acceptance of their target claims, but is also (as I said in Part I, above) concerned with the semantics of their target discourse being representational. But that semantic concern is not to the point in much which follows here, so it is left implicit for ease of exposition except where necessary.

evaluations or the evaluative, any more than 'revolutionary fictionalism' is fictionalism about revolutions, or about the property of being revolutionary.)

An evaluative fictionalist might be a hermeneutic fictionalist and might not be. Perhaps the evaluative fictionalist takes no view on hermeneutic fictionalism. Similarly, even though principled revolutionary fictionalism plausibly entails evaluative fictionalism, evaluative fictionalism does not entail revolutionary fictionalism: an evaluative fictionalist might defend the merits of nondoxastic acceptance (for a target domain) without being convinced that those merits are sufficient, all things considered, to recommend a change from current (doxastic) practice. (Perhaps the merits are just insufficient to justify such a recommendation once the incidental costs of carrying out such a recommendation are considered, for example. Or perhaps the proposal *cannot* be carried out, and the evaluative fictionalist knows this and is not prepared to recommend what cannot in fact be done, content to leave the nondoxastic model as 'just a nice idea'.⁴) So an evaluative fictionalist might be a revolutionary fictionalist and might not be. Perhaps the evaluative fictionalist takes no view on revolutionary fictionalism. (Another, less interesting, reason why evaluative fictionalism entails no commitment to revolutionary fictionalism is just that revolutionary fictionalism is *inter alia* the denial of hermeneutic fictionalism, and an evaluative fictionalist might be a hermeneutic fictionalist.) We need, then, a three-category taxonomy of fictionalist views.⁵

This chapter is an exercise in evaluative fictionalism: its purpose is to argue for one merit of nondoxastic acceptance of ordinary moral claims. As it

⁴ We ought not to be dismissive of ideas which are 'just nice ideas': of such things *ideals* are made, approximation to which can be a crucial constraint on theory about what ought to *actually* be done.

⁵ It is not necessary to treat such a three-category taxonomy as a three way *partition* of the logical space, since as we have seen evaluative fictionalism overlaps both revolutionary fictionalism (necessarily) and hermeneutic fictionalism (contingently). But taxonomies don't need to be restricted to describing partitions in the logical space.

happens, I think there are good reasons to be a hermeneutic fictionalist, and good reasons to be a revolutionary fictionalist, on the understanding of fictionalism sketched in section. But that is not to the point here. Nor is the question whether there are any *other* merits than the one I shall argue for to nondoxastic acceptance of ordinary moral claims. That question does indeed fall within the purview of evaluative fictionalism, but my intention here is not to exhaustively explore the prospects for evaluative fictionalism; it is rather to propose one argument for that view.

I hope to do enough to show that anyone who accepts the fictionalist hermeneutic claim but proposes a revolutionary proposal to the effect that we would be better off *believing* the moral claims we accept – or who doesn't accept the hermeneutic claim but nonetheless proposes a *counter-revolutionary* proposal to the effect that if we ever went about moral deliberation on the fictionalist model we ought to seek to rectify that 'mistake' – owes an account of how the advantages of (implementing) their doxastic model outweigh the advantages of (implementing) the model I have discussed, and that it is not enough to just *assume* that a doxastic model of ordinary moral judgement affords the most appealing vision of our ordinary moral practice. Put another way, the purpose of this chapter is to dislodge any presumption that nondoxastic ordinary moral judgement and deliberation would have to be, at best, a necessary *compromise*.

The sense in which I want to say that a nondoxastic model is attractive here is not the sense in which the *model* itself is attractive (*qua* model, for theoretical purposes), but in which *things being the way the model represents* is attractive. That is, I shall argue not for the theoretical virtues of fictionalism (the virtues of those positions as theories), but for the practical advantages of nondoxastic acceptance of ordinary moral claims.

There is a worry which might be raised against the idea that evaluative *moral* fictionalism could be right, a worry about the relation of morality to

practical reason. According to that worry, there could be nothing *good* about a practice of practical *reason* which abdicated judgement and deliberation to some nondoxastic attitude, for that very abdication must amount to the renunciation of just the sorts of attitudes (moral beliefs) the *rational* normative governance of which affords the only hope of preserving a link between moral attitudes and anything deserving of the name 'rationality'.

But the role of the nondoxastic is not limited to fulfilling functions outside of the domain of activity governed by reasons, and nondoxastic commitments are not, thereby at least, precluded from playing important and rationally respectable roles in the determination of actions or the forming of judgements. That is, we ought not to think that in affording the nondoxastic a role in ordinary moral deliberation we are thereby retreating from the appealing and plausible idea that ordinary moral deliberation should amount, in some way or another, to a species of practical reason.

The general point is well summed up by Joseph Raz:

The value of our rational capacity, i.e., our capacity to form a view of our situation in the world and to act in light of it, derives from the fact that there are reasons that we should satisfy, and that this capacity enables us to do so. It is not, however, our only way of conforming to reasons. We are, e.g., hardwired to be alert to certain dangers and react to them instinctively and without deliberation, as we react to fire or to sudden movement in our immediate vicinity. In other contexts we do better to follow our emotions than to reason our way to action. These examples suggest that the primary value of our general ability to act by our own judgment derives from the concern to conform to reasons, and that concern can be met in a variety of ways. It is not, therefore, surprising that we find it met also in ways that come closer to obeying authority, such as making vows, taking advice, binding oneself to others long before the time for action with a promise to act in certain ways, or relying on technical devices to "take decisions for us," as when setting alarm clocks, speed limiters, etc. (Raz, 2006: p1017)

Raz's concern here is with whether it must amount to an abdication of our practical reason – our capacity, indeed our obligation *qua* rational persons, to decide for ourselves what to do on the basis of reasons – to submit to the

authority of another, and his conclusion is that so submitting ourselves is not bound to amount to such an abdication, and that in fact it can be squarely in the interests of practical reason to submit to an authority in the interests of better achieving that which we have reason to want to achieve, if the authority in question is better placed than us to know what achieving that requires. But the point, I take it, is a more general one: practical reason, as the passage from Raz points out *en route* to making that point, is not just sometimes better served by submitting to some external authority in the form of another person or a political or legal institution; it is just as likely to be sometimes better served by allowing ourselves to be subject to some 'external determination within us', a reflex, habit or unreflective mechanism external to our practical reason (that capacity which, *qua* rational creatures, makes us *persons*) but nonetheless part of our own psychology or physiology. What is going on in all these cases is that a reason we have for *deciding* (in the full deliberative sense) on some course of action is actually better served by our not deciding (in that full deliberative sense) but coming to intend or perform that action some other way.

What does this mean for the evaluative standing of the nondoxastic type of moral commitment I have proposed? It means, I think, that to the extent that moral commitments are apt to play a role in practical rationality, we ought to remain open to the possibility that those commitments are not *themselves* subject to the sorts of rational constraints which govern beliefs, just as neither bodily reflexes nor alarm clocks are subject to *rational* constraints. (Neither a reflex to move away from a fire nor an alarm clock going off at 7.00am, for examples, are rational processes, and so nor are they criticisable *as* rational processes, by their failing to conform to rational norms or requirements, for obviously none apply.) We might say that the nondoxastic moral case is something like the case of authority as Raz sees it, depending for its normative weight on the fact that it serves practical reason

by achieving things which we have reasons to want to achieve and that those reasons are the ones about which we are typically concerned when thinking about whether to submit ourselves to what might be called the 'authority of our nondoxastic commitments'. Or we might prefer to say that the reasons which legitimate our submission to the authority of our nondoxastic commitments are typically opaque to us. Either way, what is required (of the evaluative fictionalist) is to specify some reason(s) for not deliberating about morality in a way that is tied to (doxastic) rational constraints. The point of invoking Raz's observation about practical reason here is that in giving reasons of either type for not deliberating in that way we need not be saying that there are reasons for adopting moral commitments which are contrary to practical reason, any more than it is contrary to practical reason to rely on instincts for some purposes, or on alarm clocks.

In what follows I shall not tend to put things in terms of practical reason, or at least I shall make nothing of the idea that morality is more closely tied to practical reason than other domains of commitment (I shall, though, be saying a great deal about the *role* of moral commitments in practical reason). The point of these comments has been to suggest that it would be a premature and implausible objection to the evaluative fictionalist position I am about to develop to argue that the form of moral deliberation I have proposed cannot be a *good* form of moral deliberation because it fails to be in the business of responsiveness to reasons – normative facts, considerations which count in favour – at all. That objection would require a view of responsiveness to reasons which, implausibly and naively, counts only beliefs as normatively respectable responses to the reasons that there are.

4. The Subjective Warrant Argument

David Papineau's version of what I call a 'London argument' for fictionalism⁶ invokes the fact that 'if moral [...] judgements were beliefs, then the bridge principles which take us from natural premises to moral [...] conclusions would be unwarranted' and that 'it is only when we view moral [...] conclusions as non-doxastic that we are free to read those bridge principles as acceptable prescriptions, rather than unevidenced beliefs', concluding that 'we need to deny the title of belief to moral [...] claims, if we are to continue upholding them' (Papineau, 1993: ch6, sec.13, pp201-2).⁷ There is, here, the germ of a good argument for evaluative fictionalism, in my sense of fictionalism, and I shall say more about how to make the argument good for that purpose below; but the way Papineau is thinking of the argument is not good for that purpose for a number of reasons.

For one thing, it is pretty clear that Papineau is not committed to any version of *semantic representationalism*: he casts the nondoxastic alternative as 'the option of reading moral judgements as expressing some kind of impartial approval', and lest we are tempted to read this as just a commitment to some sort of representational subjectivism (taking moral

⁶ A 'London argument' for fictionalism is an argument explicitly appealing to the norms accepted as governing the target practice and discourse. Such arguments are given by Papineau (1993), Kalderon (2005a) and I (in the previous chapter), all of whom do our philosophy in London. Non-London philosophers have tended to offer other sorts of arguments for fictionalism, so I take the liberty of calling our sorts of arguments 'London arguments'. Incidentally, care is required in characterising precisely what the role of the appealed to norms *is* in assessing these arguments – Sainsbury (2009: ch9) criticises Kalderon, 2005a (ch1) for saying that we are not subject to particular (epistemic) norms with respect to our moral commitments when, in fact, we are; but in fact Sainsbury's criticism is a *non-sequitor*, since Kalderon's argument (on behalf of the hermeneutic fictionalist) is actually predicated on a claim about which norms we are (typically) *sensitive to*, whether or not, as Sainsbury claims, they obtain anyway.

⁷ Papineau runs precisely the same nondoxasticism argument for moral and modal claims, and indeed the omissions from the quoted text are just where Papineau says 'or modal'. I have no interest here in discussing whether modal claims are or ought to be accepted nondoxastically, so I have left out mention of modality to avoid unnecessary distractions. Certainly Papineau's argument does not rely on the nondoxastic acceptance of modal claims to ground the nondoxastic acceptance of moral claims (they are separable, parallel arguments), so no harm can come of this for our purposes here.

claims to *describe* psychological states, or relations of psychological states to actions or states of affairs) he adds in a footnote that this is ‘as, for example, in Ayer’, citing chapter 6 of *Language Truth and Logic*, a *locus classicus* (if not the best example) of an emotivist position which denies semantic representationalism in having the semantics of moral claims depend upon their *evincing*, not describing, psychological states (Papineau, 1993: 199-200). Indeed, his talk of ‘acceptable *prescriptions*’ in the argument quoted above is also reminiscent of Hare’s prescriptivism (in, for example, Hare, 1952), also a theory on which the full meaning of moral claims cannot be reduced to their descriptive content. And sometimes (Papineau, 1993: p201, n18) he seems to be treating Blackburn’s expressivism (in, for example, Blackburn: 1984; [1986]) as a version of the idea he has in mind.⁸ The point about all this, for our purposes, is just that Papineau is clearly not thinking that one of the conditions for a fictionalist analysis in my sense is satisfied by the picture of moral commitment he is proposing, for that picture is, if his passing comments glossing his idea in terms of these types of theories are to be trusted, an *anti-representationalist* picture.

Nonetheless, the argument he sketches may still be useful for our purposes. All we must do to fix it in this respect is to avoid those anti-representationalist types of glosses, and that is not hard. Let us, in fact, just

⁸ This is not really the place to worry about what Papineau’s commitments concerning the possibilities here are, but it was certainly Hare’s view that Ayer’s emotivism (in Ayer, [1936/46]) and his own prescriptivism are distinct in respect of both their commitments and their truth, so if he was right (about either of those things) then there is some tension in Papineau’s apparent flirtation with both emotivism *and* prescriptivism, never mind the role of more sophisticated forms of expressivism such as Blackburn’s. (Hare’s own reasons for thinking that emotivism is essentially different from his own prescriptivism and inferior to it (which I report rather than endorse) were various, ranging from the thought that emotivism seems to rely on the implausible idea that we are in possession of psychological state concepts at the early stage of development at which we grasp the meaning of sentences in the imperative mood (Hare, 1952: ch1 sec6) to the thought that emotivism relies on the *perlocutionary* aspects of moral language, of which there can be no *logic* and hence no account of the logic of moral commitment, whilst his own prescriptivism relies on the *illocutionary* aspects of moral language, of which there is a logic (see e.g. Hare, 1997: ch6).)

assume that the semantics of moral discourse are purely descriptive in the ordinary way and run the argument just the same.

Once we have expunged the strictly extraneous denial of semantic representationalism, the argument, filled out a little more but still not the argument I want to endorse, looks like this. We have good reason not to be sceptical about the thought that at least some of our moral commitments are fit to be upheld. But the warrant for beliefs depends upon the good evidence there is for their truth or probable truth. And there could be no good evidence for moral commitments, so if they are beliefs they must not be upheld, for they could thereby not be warranted. So moral commitments had better be nondoxastic, for nondoxastic commitments are fit to be upheld even in the absence of epistemic warrant.

That argument, which I'll call the 'Objective Warrant Argument' turns, of course, on a premise to which a sceptic about (knowable) moral facts or truths, even if not a sceptic about the value of upholding at least some of our moral commitments, is entitled, at least on some not implausible views about the relation of good evidence to what is true. But of course a realist is not liable to accept that premise, and not even all non-realists are entitled to accept it, at least if they think of epistemic warrant so as to allow that there can be warrant for beliefs which are not true. The question most obviously begged by the Objective Warrant Argument is whether there *is* absolute warrant for moral beliefs.

The weakness of this argument is linked to an issue we have come across before. In discussing the nature of various modes of acceptance (in Chapter Three, above), I spoke of a person's own sensitivity to conditions which are seen or thought (by them) to bear on the correctness of a claim up for acceptance. And the version of the Warrant Argument I want to defend replaces the controversial (and dubious) premise about *objective* warrant with

a set of premises about *subjective* or *perceived* warrant, and the role of subjective or perceived warrant in relation to action guiding.

Concern for action guiding is not, I think, entirely absent from the ideas behind the Objective Warrant Argument, for it is plausible that the interest we have in upholding (some) moral commitments is at least partly to do with the way in which upholding (some) moral commitments is practically necessary for doing some of the sorts of things we think it is good to do. (But see my discussion of this thought in the next chapter.) The Subjective Warrant Argument, however, takes that way of motivating the premise about the value of upholding (some) moral commitments and links that interest in action guiding with the action guiding consequences of what I want to call lack of subjective epistemic warrant.

The Strong Subjective Warrant Argument

The (Strong) Subjective Warrant Argument, then, is as follows.

(Strong) Subjective Warrant Argument

We have good reason to think that at least some of our moral commitments ought to be fully upheld, at least in part because only fully upheld commitments are apt to play action guiding roles in important contexts. But we ought not to fully uphold *beliefs* for which we lack *subjective (epistemic) warrant*. Subjective warrant for *beliefs* depends upon the good *epistemic* reason a person thinks they have for believing the claim they believe. And there could be no reliable way of ensuring that a person thinks they have good *epistemic* reasons for their *moral* commitments, so if those commitments are beliefs there could be no reliable way of ensuring that they be fully upheld, for they could thereby not be subjectively warranted, ought therefore to be less than

fully upheld, and thus will not be fully upheld by those sensitive to the subjective warrant norm governing beliefs. Since beliefs which are not fully upheld are ill-suited to play action guiding roles in important contexts, moral commitments had better be *nondoxastic*, for nondoxastic commitments ought to be fully upheld, and hence are apt to play action guiding roles in important contexts (for those sensitive to the norms governing the upholding of those commitments), even in the absence of subjective epistemic warrant, and are thus liable to actually play those roles even in the absence of subjective epistemic warrant.

That, essentially, is the argument. But it needs a great deal of unpacking. The first thing is to be clear what is meant by *subjective warrant*. Perhaps the most important thing to understand about what I am calling subjective (epistemic) warrant is that it is *not* credence. The way I am thinking about beliefs is not committed to making room for anything like the Bayesian notion of a degree of belief at all (nor is it strongly committed to ruling it out, for our purposes here at least). I am quite prepared to countenance degrees of subjective warrant; but I insist that degrees of subjective warrant are not (or would not be) analogous with credence, on a Bayesian notion of credence. What I am calling subjective (epistemic) warrant is a person's (reasonable) confidence in the *grounds* for their belief, or in their belief having good grounds. Notice that a person might *fully* believe that *p*, but not have much confidence in their grounds for that belief; and that a person *might* (notwithstanding certain views about the 'transparency' of belief) be very confident that there are good grounds for believing that *p*, but not believe it at all. I take it that coming to think of one's belief as ungrounded (coming to lack subjective warrant) is *not* sufficient to undermine the *force* of that belief.⁹ Indeed, if

⁹ This is a point which, I think, Hawthorne, 2004 is tentatively endorsing (2004: p169), though he is open to the idea that 'perhaps some case could be made that outright belief is

something like Harman's (1986: ch6) 'conservativeness' norm for beliefs is right it is not only possible but also quite right to maintain a belief until one has good evidence that it is *false*, not merely that it is ungrounded. (Ungrounded beliefs can, of course, be true nonetheless.)

Which brings us on nicely to the issue of 'fully' upheld beliefs (and commitments in general). If one can continue to fully believe something which one lacks subjective warrant for, by lacking confidence in one's grounds for that belief, then what is lacking in the way of 'upholding' such that that belief is not apt to play its full action guiding role (as is required for the Subjective Warrant Argument)? The answer is that 'upholding' a belief in the sense intended here is a matter of giving it a place (a role) in one's judgmental and deliberative economy, allowing it to be used in inferences, for example, or employing it as the basis for one's decisions. A fully upheld belief is a belief that is allowed to play all the roles that beliefs are normally apt to play. Less than fully upheld beliefs are beliefs upon which one places unusual restrictions in respect of the roles they are allowed to play, or the contexts in which they are allowed to play them.

So, for example, suppose I find myself believing that I have enough petrol in the tank to get home on the motorway without filling up at the approaching petrol station. But I also know that the fuel gauge is broken, and that my unreliable friend borrowed my car yesterday and might have

often destroyed by the salient possibility of error even though subjective confidence is not' (p170, n24).

I want to flag up here that I am not yet quite sure what I think about the relation of confidence in the grounds of a belief to credence as confidence *in* a belief (though I am sure that I reject the Bayesian reduction of belief to credence, for it seems obvious that a person's belief and a person's confidence in their belief are two distinct things, and it is unclear to me why the Bayesian is required to reduce the former to the latter in order to entitle themselves to any of the results they have good reason to want – treating their calculus as a calculus of *confidence* seemingly allows for all the claims they want to make about rationality whether or not it is allowed that belief is non-reducible to confidence above a certain threshold.) I don't think anything I have to say here, though, depends upon deciding what that relation between confidence in *grounds* and credence is.

done a lot of driving without bothering to refill the tank. So, reflecting on my belief that there is enough petrol in the tank, I might very well decide that I have no good grounds for that belief even though I cannot shake it and continue to believe that I have no need to stop at the approaching petrol station. Suppose that, on the basis of my lack of subjective warrant, I think that even though I believe I've got enough petrol (perhaps I would be *surprised*, despite my better judgement, if it turns out that I don't have enough), I had better not base my decision whether to stop at the approaching petrol station on that belief, precisely because *that* belief is one of the ones I lack the right sort of confidence in the grounds of, and this is a context in which I had better rely, to the best of my ability, on *well grounded* beliefs since being wrong really matters (I very much don't want to run out of petrol on the motorway). Then my belief that I have enough petrol to get home on the motorway without stopping at the approaching petrol station is not being *fully* upheld.

(If you think that I can simultaneously believe that p and believe that *not- p* , then here's another case. Suppose I find myself believing that p and believing that *not- p* . I might think (for good classical logic reasons) that I must thereby have at least one false belief, and because I have as good grounds for one belief as for the other, I might reason that my grounds for *each* belief cannot be relied upon (for the grounds for one of them must be bad, and I don't know which the bad one is). I still believe each, for there are *apparently* grounds for each, which makes each compelling. But there will be *some* deliberative contexts in which my doubts about the grounds of each will mean that I will be tentative about relying on them in inferences. Again, those beliefs are thereby not fully upheld, though this time the explanation is to do with a judgement about how good one's apparent grounds are likely to be, not with one's awareness that one has no grounds (nor even apparent ones).)

It might seem that this discussion is covering ground already familiar from the contemporary debate in epistemology about ‘contextualism’, a debate (or cluster of debates) about whether knowledge (or the verb ‘know’) is context dependent, and whether the ‘stakes’ involved in various contexts of knowledge attribution (including to oneself) play a role in determining whether someone counts as knowing. But in fact the point I want to make here about *belief* and its role in theoretical and, in particular, practical deliberation is intended to be completely silent on anything to do with *knowledge*, and I take it that I am not in the business here of agreeing or disagreeing with anything that contextualists about knowledge, and their critics, have to say.

The Knowledge Norm of Action and Practical Deliberation

What, though, of those (such as Hawthorne, 2004 and Hawthorne & Stanley, 2008) who think that the norms of practical deliberation are *knowledge* norms? If they are right, isn’t it the case that my discussion of belief, leaving knowledge aside, is misguided, because our readiness to rely on beliefs in practical deliberation ought to depend upon whether we take ourselves to *know* a believed proposition, and not upon our attitudes to the grounds of our beliefs just *qua* beliefs?

This is not the occasion on which to go in for a proper discussion of putative knowledge norms governing action. Suffice it to say, for now, that I am sceptical (of course I’m not the only one: Neta, 2009 has some reasonably *prima facie* plausible counterexamples, and an alternative diagnosis of what is going on in Hawthorne & Stanley’s cases; see Brown, 2008 for a different attack on the knowledge norm.¹⁰). But be the actual norms (objectively)

¹⁰ I leave aside, for reasons of space, discussion of Williamson’s view that though the knowledge norm defended by Hawthorne and Stanley is not right, the permissibility of our

governing practical deliberation as they may, it will be possible to run a version of the Subjective Warrant Argument which appeals to our tendency, if we are epistemically responsible, to not rely on beliefs which we take to be ungrounded in important contexts anyway. That version of the argument replaces the talk of what a person sensitive to the norms that there are is liable to do, in the argument as presented above, with talk of what we are bound to do in the interests of respecting even a knowledge norm of practical deliberation given the failure of the KK principle: since, in weighing considerations in the course of high stakes practical deliberation, I do not necessarily know which of my beliefs are known, it is fitting for me to rely only on those beliefs for which I take myself to have good grounds, on the basis that well-grounded beliefs are more likely to be knowledge than ungrounded ones. This pragmatic version of the argument, focussing on what epistemic responsibility implies about which considerations I ought to give weight to in my practical deliberation, can be run whilst granting the knowledge norm of action, so the truth of the knowledge norm thesis does not automatically undermine the spirit of the Subjective Warrant Argument.

But does a challenge from those who maintain that knowledge is the norm of reliance on claims for practical deliberation not remain? The challenge might be this. Nondoxastic attitudes are (presumably) not apt to constitute knowledge as beliefs are, so whilst it might be the case that the perceived grounds of our own beliefs is relevant in the epistemological way just suggested (relevant, that is, as a *guide* to which beliefs we ought to draw practical inferences from or act on according to the knowledge norm), nondoxastic attitudes must necessarily be just as badly off with respect to

employing some claim in our practical reasoning is sufficient for our knowing it (Williamson, 2005), a view which though it rejects the knowledge norm raises similar issues for the Subjective Warrant Argument. Suffice it to say for now that Williamson's view is no *more* problematic than the knowledge norm view, for it is logically weaker, so if what I go on to say about the relevance of the putative knowledge norm is right then Williamson's view is similarly dealt with.

our taking them as likely to be knowledge as our most seemingly ungrounded beliefs are. So even if it is the case that, if we are sensitive to the demands of epistemic responsibility and the knowledge norm of practical deliberation, we would not act on the basis of our seemingly ungrounded beliefs, nor is it the case that we would act on *any* of our nondoxastic attitudes. Thus, the nondoxastic picture is a hopelessly dystopian one on which we typically *never* act on our moral commitments in high stakes contexts, if we are sensitive to the norms of practical deliberation and action. And such a dystopian picture would, far from delivering the result the evaluative fictionalist wants, put nondoxasticism in a far worse position in respect of the benefits of things being *its* way than doxasticism could ever be.

There might be something to this challenge, but for it to have real bite against the Subjective Warrant Argument it would have to be granted that the only dimension along which to assess the claims we accept as the basis of rational action is the epistemic dimension for which knowledge is said to be the relevant measure. Put another way, it would have to be granted that the knowledge norm applies to any and all of our attitudes involved in practical deliberation, for otherwise it would be open to us to grant the knowledge norm for those *doxastic* attitudes involved in our practical deliberation but say that some other norm(s) governs those *nondoxastic* attitudes involved in our practical deliberation, and that therefore the fact (if it is a fact) that no nondoxastic attitude could ever be suspected of amounting to knowledge would not (*contra* the challenge under consideration) mean that no nondoxastic attitude could ever satisfy the norms governing reliance on some claim for practical deliberation. And in fact it is very plausible that the epistemic dimension is *not* the only dimension along which to assess *all* of our attitudes which play practical deliberative roles.

If you think that *desires*, for example, play such roles (and play them quite properly) then you had better deny that no attitude which is not apt to

amount to knowledge could be a permissible basis for practical inference or action.¹¹ And if there are some nondoxastic attitudes apt to play such roles (and play them quite properly), what reason is there to think that the nondoxastic attitude the fictionalist model invokes is not such an attitude? In general, there is no reason to expect that nondoxastic attitudes, just in virtue of their being nondoxastic, are bound to fail to satisfy the norms governing practical deliberation whatever they are. Rather, the fact that nondoxastic attitudes do play permissible roles in practical deliberation, with full action-guiding force, is reason to expect that there are norms governing those attitudes which they are apt to satisfy. So if the challenge under consideration requires that there are no such norms (that the only norm is a knowledge norm which nondoxastic attitudes are bound to fail to meet), we had better reject the challenge on that very basis.

Whether or not those who think that there is a knowledge norm of practical deliberation or action are committed to that being the *only* norm governing the attitudes which play practical deliberative and action guiding roles is a moot point. They might be content to say that so far as attitudes are *doxastic* they ought only to be relied upon in practical inferences and as the basis for action if they amount to knowledge, but that what determines the permissibility of relying on other, *nondoxastic* attitudes is, for all they want

¹¹ Perhaps you don't think that desires do ever play a *deliberative* role, and that our practical deliberation is a matter of our reasoning from our beliefs about our desires rather than from those desires themselves. But if you were to think that, why stop there? Why not invoke beliefs about our beliefs to play the actual deliberative roles we think ordinary first-order beliefs play? If it is necessary to invoke a *belief about* my desire for a drink to explain my rational decision to go to the kitchen, why is it not also necessary to invoke a *belief about* my belief that there is a drink in the kitchen? The point is that the only vaguely plausible reason to invoke beliefs about desires rather than desires themselves is to respect the thought that it is *awareness of* a desire (and not just the *having* of a desire) which gives me a reason to act. But if you think that the way to respect that thought (on the assumption that it deserves respecting) is to invoke a *belief about* a desire, then presumably respecting the thought that our practical *reasoning* must rely on beliefs we are *aware* of must likewise invoke second order beliefs to account for that awareness. Which, I take it, amounts to an informal *reductio* of the assumption that the awareness requirement entails a belief requirement.

to say, an open question. If they say that, which seems like the most plausible thing for them to do if they grant that quite proper practical deliberation can rely on such things as desires, or that such things as desires can be good reasons for actions, and that we can quite properly treat them as such, then the mooted challenge to the Subjective Warrant Argument which we have been concerned with is a non-starter.

The defender of the knowledge norm might say that desires are not going to do the job I suggested, namely the job of exemplifying how attitudes which play a quite proper role in our practical deliberation and action are in some cases subject to some norm other than the knowledge norm, because desires are not normatively governed in practical deliberation and with respect to action at all: desires are had or not had, and whether they *ought* to be had or not had is not a question which arises, so adverting to desires in the interests of motivating the thought that there is some non-epistemic *normative* aspect of our practical deliberation and rational action is hopeless. But that way of resisting the point I am pressing is both open to serious challenge itself and, more importantly, beside the point.

It is open to serious challenge because it is by no means obvious that desires are *not* normatively governed. Certainly, the thought that a particular desire or aversion might be *appropriate* or *inappropriate*, or that it might be *praiseworthy* or *blameworthy*, or indeed that it is properly the object of various other normative or normativity-related appraisals is not an unfamiliar one. (See Gibbard, 1990 for just one sustained philosophical working out of that thought.)

More importantly, resisting the point that the role of desires seems to show that there is normative governance of some attitudes which play a role in our practical deliberation and rational action which is not epistemic (so certainly not the knowledge norm) by arguing that desires are not normatively governed themselves is beside the point. The point is that their

role in practical deliberation and rational action plausibly *is* normatively governed, even if whether one has or lacks a particular desire or aversion is not. That is to say that once we have a desire, the question whether or not we *ought* to have that desire is irrelevant to the question what role that desire ought to play in our practical deliberation and rational action (so it is also irrelevant if the question whether we ought to have that desire or not has no answer, or doesn't arise).

Finally (for now), we might want to know what, in the absence of reasons to *rule out* nondoxastic attitudes such as the attitudes invoked in the fictionalist's account of moral commitment from being normatively governed by non-epistemic norms (such as the knowledge norm) and satisfying those norms, the reasons are for expecting that those nondoxastic attitudes *will* be so governed. Of course the important issue with respect to the Subjective Warrant Argument is not immediately whether they are so governed; it is whether those sensitive to the norms governing practical deliberation and action are likely to act on such nondoxastic attitudes in high stakes situations.¹² But it will be easier to believe that those sensitive to the norms governing practical deliberation and action *will* act on their nondoxastic commitments if we have some idea what the norm(s) to which they are sensitive in so acting is.

I shall make only a sketchy suggestion here, to get the ball rolling. Call the norm to which our nondoxastic moral commitments are subject in

¹² At this point it is worth briefly saying something about why I am invoking *sensitivity* here. We might expect, given the emphasis I have been placing upon our own perceptions of the normative standing of our commitments and the possible divergence of that from the *actual* normative standing of those commitments, that I would be content to show that we have reason to expect that we just will act on our nondoxastic commitments in high stakes circumstances for *whatever* reason. And it is true that so much would suffice for the Subjective Warrant Argument. But unless we think that *rational*, or at least *normatively approved*, action and deliberation about what to do is a rarity, it makes for a more reasonable claim that we will act on our nondoxastic commitments in high stakes circumstances if we can see that that is because there is *reason* to so act, and that we are typically sensitive to such reasons.

respect of their role in practical deliberation and action 'Integrity'. Roughly, Integrity requires that a person act on a moral commitment or rely on that commitment in practical deliberation only if that commitment is a commitment which they believe satisfies the correctness condition (to do with accepting some moral claim being required for being the sort of person with the virtues one values) mentioned in the Character-Centred model (described in Chapter Four). That is *not* to say, of course, that Integrity requires that we believe the moral claims which we are committed to if we are to rely on them in exercising our practical reason: what (and *all* that) is required by Integrity in the way of belief is that we believe that *acceptance* of a claim upon which we are to rely really is required for manifesting the virtue we seek to manifest, and *that* says nothing at all about what the mode of acceptance thereby invoked is.

Integrity is really just a (rough) codification of the thought that we ought to act on and deliberate from moral commitments which we take to be correct, and ought not to act on and deliberate from moral commitments which we do not take to be correct. The particular way in which Integrity encodes that thought is driven by the idea, which is a feature of the fictionalist picture I have sketched, that we might typically take our moral commitments to be correct on the basis of their satisfying some normative condition that is not truth-associated and is rather associated with the connection between our moral attitudes and our character, in which we might typically take a great deal of interest. Thus, Integrity is supposed to capture the sense in which our actions and deliberations about what to do are, in respect of their involving *moral* commitments, a matter of doing or resolving to do what we expect of ourselves as, for example, kind, or honest, or God-fearing people. (And our reasons for wanting to be kind, or honest or God-fearing are, on the Character-Centred picture, not to do with those virtues being *epistemically* desirable – a matter of being someone who *knows*

that helping people is the thing to do because it is kind, for example – but with their being appealing as a way to be for whatever reason: a person might prefer to be a strictly deluded philanthrope than a knowledgeable miser – they might be sanguine about the possibility of the moral *truth* being that generosity is not required, regarding that truth, if it is a truth, as at best uninteresting and at worst dangerous).

So the Character-Centred picture plus a conception of the normativity of practical moral deliberation and action which invokes Integrity promises to secure a sense in which our nondoxastic moral commitments are normatively regulated, and regulated by norms which there is reason to think our nondoxastic moral commitments can and often do satisfy. There is, then, at least *prima facie* reason to expect that the typical moral deliberations and actions of those who are typically sensitive to the norms governing practical reason will *not* be undermined by lack of *epistemic* subjective warrant for particular moral claims, acceptance of which claims *will* be fully upheld due to their managing to satisfy the *non-epistemic* norm (Integrity) governing them.¹³ I am not, certainly on the basis of what I have said here, convinced that Integrity, even as sketchily worked out as here, is in fact the norm of moral practical deliberation and action, and I do not expect to have convinced anyone else. That is work for another occasion. What I have tried to do here is just to point out that we can think of analogues of the knowledge norm for practical beliefs (or whatever epistemic norm governs

¹³ It is probably not ideal to say that Integrity is a ‘non-epistemic’ norm, since Integrity does demand that we rely only on those accepted claims which we *believe* satisfy the appropriate correctness condition. Nonetheless, since the relevant correctness condition, on the Character-Centred View, is itself non-epistemic, it is convenient and not *too* misleading (I hope) to say that Integrity is non-epistemic: certainly it is not epistemic in the same sense as the knowledge norm is epistemic, for the knowledge norm ‘bottoms out’ in epistemic concepts in a way Integrity on the Character-Centred View, which bottoms out by referring to a correctness condition that has nothing to do with truth, evidence, belief or anything similar, does not. There is much in common here, I think, with van Fraassen, 1980 on the idea that beliefs about the empirical adequacy of particular theories are apt to replace belief *in* particular theories for the purposes of rational theory construction.

practical beliefs) for practical nondoxastic moral commitments, thereby rebutting the worry that the nondoxastic must spell disaster for our hopes of maintaining a genuinely *normative* conception of practical reason.

Defending the Main Premise of the Strong Argument

I take it that cases such as the motorway petrol refill case (above) make it plausible that lack of subjective warrant *is* sufficient to undermine our fully upholding – in the sense of giving full deliberative weight to – a belief. The point is that *if* a person lacks subjective epistemic warrant for a belief, then they are liable to (reasonably) prescind from acting on that belief. The same cannot be said for the variety of moral commitment I have described according to the Character-Centred model in Chapter Four. It is *not* the case that lack of subjective *epistemic* warrant is liable to undermine the action guiding (or other practical deliberative) roles of commitments we treat as required by the demands of being a person with the virtues we value. That is why, as the Subjective Warrant Argument says, we have reason to prefer our moral commitments to be as my view says they are than to be beliefs, *if* moral beliefs would be particularly subject to lack of subjective epistemic warrant.

But *are*, or *would*, our ordinary moral beliefs be subject to lack of subjective epistemic warrant? That is the crucial move in the argument, and it is clearly the analogue of the problematic premise in the Objective Warrant Argument. Resources other than those appropriate for deciding whether the lack of objective warrant premise is true are required for deciding whether the lack of subjective warrant premise is true: the lack of objective warrant premise required some contentious views about, at least, the nature of objective epistemic warrant and, depending on which views one has about that, some strong metaphysical commitments (namely, to non-realism about

moral facts); but the lack of subjective warrant premise requires no such commitments, instead requiring some commitments about our typical deliberative practices. In particular, what is required to defend the lack of subjective epistemic warrant premise is some evidence which suggests that our subjective epistemic warrant for moral beliefs is not particularly secure: we need to be assured that there is something about our typical thought about the grounds of our moral beliefs which does or would undermine subjective epistemic warrant.

What is *not* required is some reason to think that we *never* have or would have subjective epistemic warrant for moral beliefs. That our resolve in acting on moral beliefs is typically undermined by lack of subjective epistemic warrant is a reason to prefer our moral commitment to typically be of the sort I have described if there is a *general* tendency of people to lack subjective epistemic warrant. (This is just what putting the point in terms of the *reliability* or *security* of our subjective epistemic warrant is supposed to achieve.) After all, there is no reasonable hope of all and only those who *have* subjective epistemic warrant for moral beliefs accepting moral claims doxastically whilst all and only those who would *lack* subjective epistemic warrant for moral beliefs accept moral claims in the way I have described. So if the tendency to lack subjective epistemic warrant for moral claims, and to therefore lack resolve in acting on moral beliefs, is even moderately general, then there is reason to prefer *typical* acceptance of moral claims to be as I have suggested: if it were, then even those who do lack the relevant subjective epistemic warrant would nonetheless be free to act on their commitment with full resolve even in hard cases (free, that is, to fully uphold their moral commitments, which we have reason to want), and those who do in fact have the relevant subjective epistemic warrant would be no less likely to act on their moral commitments with full resolve than if they believed

them, for they would have just the same freedom to do so as those who lack the relevant warrant. Call this the 'Safety Net Idea'.

All we need, then, to support the lack of subjective epistemic warrant premise is a reason to think that the tendency to think that we have no good evidence for our moral beliefs is or would be moderately general. One reason is, I think, provided by reflecting upon moral disagreements. Once again, it is crucial to understand that the point here is not about whether moral disagreements tell us about *objective* warrant, or anything about the *real* normative status of moral beliefs. We are not, therefore, in the realm of currently fashionable debates about the epistemology of disagreement between 'epistemic peers'. The point is just that moral disagreements are able to tell us something about people's *attitudes* to evidence for moral claims and the epistemic status of moral beliefs. Subjective epistemic warrant is undermined or lacking if *from the point of view of the believer* there is no reason to have confidence in their belief, though they maintain that belief; whether or not rationality or objective warrant supports that view of theirs about their reasons for confidence. And moral disagreements often seem to involve exchanges of ideas which invite the suspicion that the parties lack subjective epistemic warrant.¹⁴

Certainly it has seemed to many theorists, such as Stevenson, a more careful and interesting emotivist than Ayer, that typical disagreements about morality are something other than exchanges of reasons for moral beliefs. Indeed Stevenson (1944) starts out his investigation of the meaning of moral language with just such a thought (see also e.g. Stevenson, [1948]). The mistake of the emotivists, including Stevenson, was to think that the fact that something other than exchanges of reasons for moral beliefs is at issue in moral disagreements leads to needing to treat the *content* of what is

¹⁴ For an up-to-date collection of discussions about the relevance of disagreement to the epistemology of *objective* warrant or rationality, see Feldman & Warfield, 2010.

disagreed about non-representationally. There are, in fact, two other options: we might say (rather implausibly, but not obviously hopelessly) that what is going on in such disagreements is disagreement about beliefs, but that the practice of arguing about them does not amount to the giving of reasons for those beliefs; or we might say (rather more plausibly, I think) that though Stevenson and his ilk are right to call such disagreements 'disagreements in attitude' and to treat such disagreements as different from 'disagreements in belief', nothing follows from *that* with respect to the content of what is argued about, for we can have nondoxastic commitments with fully representational content and might be disagreeing about the propriety of those nondoxastic commitments.

What seems right in Stevenson's observation, though, is that whatever is typically going on, it is not plausibly the serious giving of reasons for beliefs. By 'serious' here I mean the giving of reasons which is sincerely meant as persuasive in respect of explaining to another person what one's evidence for the beliefs one is defending is. Rather, as Stevenson thought, what seems more typical is the attempt to change one's discursive opponent's mind by whatever means one can. And when called upon to defend our moral commitments, we do not typically think of the reasons we have for those commitments as amounting to good *epistemic evidence* – as demonstrated by the fact that we are not at all surprised when a person we consider to be perfectly *epistemically* competent remains unmoved by being presented with that evidence. We tend to settle, in the absence of serious hopes for *demonstrating* that we are right on the basis of evidence, for agreeing to disagree. Of course, having resigned ourselves to the inevitability of disagreement, we often continue to make an effort to get certain things done, or to get people to do things. But that is typically because we remain committed to our view of what ought to be done, and are content to settle for

seeing that it *is* done, regardless of whether we can convince our opponent that it ought to be.

This phenomenon, I think, is most plausibly understood as expressing and reinforcing lack of subjective epistemic warrant. We are not typically convinced that we have good *epistemic* reason for our moral commitments, and this is reflected in our typically half-hearted attempts at adducing evidence in the hope of changing others' beliefs, and is reinforced by the experience of finding ourselves unable to convince those who we think of as (otherwise) epistemically competent: if someone is epistemically competent and presented with a good epistemic reason for a belief, then other things being equal we expect that they will form that belief; so since we take at least some people who are unmoved by our reasons for moral beliefs to be epistemically competent, we are liable to think that our reasons are not epistemically all that good, after all. (Again, the point is not that we are *right* to come to think such a thing, but just that we often *do*.) If we think for whatever reason – and our experience of moral disagreement is only one plausible reason – that our moral beliefs are not supported by *good* epistemic reasons, then we lack subjective epistemic warrant.

Amongst those who are not tempted by the emotivists' treatment of disagreement but still take the phenomenon seriously, there is a way of treating it which connects more straightforwardly with the Subjective Warrant Argument. Here, for example, is Sidgwick on disagreement about 'Common Sense' morality:

Since it is implied in the very notion of Truth that it is the same for all minds, the denial by another of a proposition that I have affirmed has a tendency to impair my confidence in its validity. [...] And it will be easily seen that the absence of such disagreement must remain an indispensable negative condition of the certainty of our beliefs. For if I find any of my judgements, intuitive or inferential, in direct conflict with a judgment of some other mind, there must be error somewhere: and if I have no more reason to suspect error in the other mind than in my own, reflective comparison between the two

judgements necessarily reduces me temporarily to a state of neutrality. And though the total result in my mind is not exactly suspense of judgement, but an alternation and conflict between positive affirmation by one act of thought and the neutrality that is the result of another, it is obviously something very different from scientific certitude. ([1874/1907]: p341-2)

This passage, which introduces what Crisp, 2007 calls the ‘Consensus Condition’, has, for some, been a jumping-off point for a discussion of the *objective* justification of our moral commitments and the sensitivity of that justification to the existence or possibility of disagreement (see, for example, Crisp, 2007 and Audi’s (2007: pp204-8) reply), and that is certainly in tune with the way in which Sidgwick seems to be thinking of the *importance* of the Consensus Condition.¹⁵ But the letter of Sidgwick’s comments chimes well with what I have been pressing here: ¹⁶ ‘the denial by another of a proposition that I have affirmed *has a tendency to impair my confidence in its*

¹⁵ Of course arguments from disagreement are a familiar feature of debates about both moral epistemology and the metaphysics of morals, regardless of Sidgwick’s contribution. See, for example, the writings collected in Gowans, 2000 for a representative selection.

¹⁶ Actually, Wedgwood, 2010 draws attention to a number of passages in which Sidgwick is explicit about the fact that he subscribes to a *normative* version of the view which I am treating here as merely descriptive of our subjective warrant opinions. Wedgwood’s discussion of what he calls ‘Sidgwick’s Principle’ is extremely interesting and offers a nice (albeit Bayesian) theory of rational belief revision on which it is very often perfectly rational to resist changing one’s view, or the strength of one’s view, in the face of disagreement with someone who *seemed* to be an epistemic peer; his point is that on some respectable construals of epistemic peerage, it is quite in order to decide *just on the basis of disagreement of moral intuitions* that a person is not your epistemic peer after all, and thereby maintain confidence in your own judgment. I have some sympathy with Wedgwood’s rejection of Sidgwick’s Principle in its normative guise; but as I hope has been made clear, I am interested here not in anything distinctive of Sidgwick’s *epistemological* views, but in a phenomenon (actual loss of subjective epistemic warrant) which he seems to *describe* quite well. And to those such as Wedgwood who accept that our moral views are distinctively *intransigent* (a conclusion for which Kalderon, 2005: ch1 gives an argument on behalf of the fictionalist), which would seem to undermine the descriptive version of what is here variously called Sidgwick’s Principle or the Consensus Condition, I would urge caution: intransigence is very often the luxury of those whose commitments are taken rather lightly, and it is not at all clear that the levels of moral intransigence commonly observed in relatively low-stakes circumstances of moral debate and decision are reflected all the way up to the highest-stakes, most *important* deliberations and decisions. Which is not to say that high-stakes decisions are never the result of intransigent commitments – they all too depressingly are; it is merely to suppose that the sort of *general* profile of lack of subjective epistemic warrant required by the Subjective Warrant Argument is not unrealistic in the domain of interest, namely the circumstances in which moral judgement, deliberation and commitment matter most.

validity', which is to say that whether my *objective* justification for what I have affirmed (what I believe) is undermined or not, my *subjective* warrant is typically undermined.¹⁷ (Note that Sidgwick speaks of 'my confidence in its [my belief's] validity', which might be interpreted as *credence* which, as noted above, subjective warrant as I am thinking of it is *not*. But there is no reason I can think of, for all Sidgwick says or is committed to by his use of the Consensus Condition, not to read the confidence he is thinking of as what I have called subjective warrant, namely confidence in the *grounds* of a belief, since we might think of the 'validity' of a belief as being a matter of its support rather than of its truth.)¹⁸

I am not drawing attention to the Sidgwickian treatment of disagreement in morals so as to appeal to its authority. The point is rather

¹⁷ As Crisp, 2007 points out, there is no reason for Sidgwick's point to be confined to the effects of *actual* disagreement – we might expect to be just as destabilised in respect of our subjective warrant for our moral beliefs by the thought that a reasonable person *might* disagree.

¹⁸ There is a strongly Pyrrhonian whiff to the idea behind the Consensus Condition: it seems to be that *the way in which the demands of morality appear different to different reasonable people* bears some undermining relation to our confidence in our moral beliefs. (See Annas & Barnes, 1985 for an excellent account of the Pyrrhonian 'modes of scepticism'. Crisp, 2007 also notes the Pyrrhonian resonances of Sidgwick's discussion of disagreement.) But again, in thinking of the idea here as Pyrrhonian, we must remember that we are not interested so far as the Subjective Warrant Argument is concerned with any sort of *scepticism*. The Pyrrhonian arguments seem to (somehow) deliver a normative verdict to the effect that the loss of confidence we are bound to suffer in the face of conflicting appearances is *justified*. This is not straightforward, since on one plausible understanding of their form, such arguments terminate in a *state of mind*, and not a normative epistemological claim at all. This makes sense in so far as it saves the Pyrrhonist from the apparent difficulty of holding that there is no knowledge of anything whilst at the same time proposing an argument with a supposedly justified conclusion – on the reading on which there is no epistemological *claim* at the end of the argument, no justified *conclusion* in the relevant sense is proposed. But if we nonetheless want to insist that *somehow* the Pyrrhonian arguments from conflicting appearances are sceptical arguments in the sense of establishing that we *ought not to* have great confidence in our beliefs, then we must insist that as well as being concerned with only our *moral* beliefs, the premise required for the Subjective Warrant Argument departs from the Pyrrhonian arguments in another way, namely that whilst those arguments seem to move from the *fact* of loss of subjective warrant (or of what looks something like subjective warrant) to *approval* of that loss, the required premise for our purposes concerns *only* the fact. It is, therefore, logically weaker than the Pyrrhonian argument, and weaker indeed than the arguments of any of the objectivist epistemological arguments from disagreement which deliver a normative verdict.

that regardless of what lessons have been drawn *from* the lack of epistemic warrant which has seemed to many to come in the train of reasonable disagreement about morals, and especially regardless of what semantic or (objectivist) epistemological lessons have been drawn from that, the threat to subjective warrant for moral beliefs from reasonable disagreement is evidently a phenomenon acknowledged by (or, less partially, posited by) many of those who have no doctrinal reason to expect to find it. Put another way, the point of invoking these theorists has been to illustrate how the threat to subjective warrant for moral beliefs is not merely a contentious and *ad hoc* commitment born of just wanting to run the Subjective Warrant Argument; it is plausibly a feature of moral beliefs, and has been taken to be plausible (implicitly, perhaps) by theorists in the course of running more or less entirely unrelated arguments. And it is worth noting that those who have taken issue with Sidgwick and with others who recognise (or, less partially, posit) the delicacy of subjective warrant for moral beliefs have tended to take issue with what they have inferred *from* that, rather than with the delicacy of subjective warrant for moral beliefs itself (see, for example, Audi, 2007).

So the (Strong) Subjective Warrant Argument requires the premise that our subjective epistemic warrant for our moral beliefs is generally apt to be unstable, for only then can we run the argument that failure of subjective warrant means that we are liable to fail to fully uphold our moral beliefs, and thereby fail to muster the fortitude to act on them in the sorts of important high stakes circumstances in which we might think the action guiding force of our moral commitments is most to be valued. I suggested that reflecting on moral disagreements gives us a reason to accept such a premise. And I have suggested that such a premise has been assumed or thought to be plausible by theorists with no apparent vested interest in the truth of a premise like that. This, I think, speaks strongly in favour of its plausibility.

There need be no suggestion that lack of subjective epistemic warrant is a universal phenomenon – recall that all the Subjective Warrant Argument requires is that lack of subjective epistemic warrant for moral beliefs is moderately general. Once we *do* lack subjective epistemic warrant, for whatever reason, we are prone to lacking the fortitude to *act* on the basis of our moral beliefs in some of the most difficult and important contexts, for we are generally prone to demur (often with good reason) from acting on the basis of beliefs of *any* type for which we lack subjective epistemic warrant in contexts where our actions have important consequences. But of course we have good reason to prefer that we and others *do* have the fortitude to act in morally difficult contexts. So, the Subjective Warrant Argument concludes, we have good reason to prefer that, in the interests of maximising the possibility of such action (the Safety Net Idea), our moral commitments are of a form not susceptible to lack of fortitude on the basis of lack of subjective epistemic warrant. And nondoxastic moral commitment on the model I have been proposing (the Character-Centred model) is apt to be action guiding because of its close connection to the norms of practical reason (as explored in Chapter Four, section 3), and is apt to avoid fortitude failure in cases of lack of subjective epistemic warrant, for it has nothing to do with modes of acceptance which are constitutively tied to *epistemic* (truth-associated) reasons. So we have good reason to prefer our moral commitments to be nondoxastic on the model I have proposed than to amount to moral beliefs. If we think that *is* how our moral commitments are, then things (in respect of this) are fine as they are. If we think that they are not, then we have a reason (perhaps not overriding) to propose reform.¹⁹

¹⁹ The Subjective Warrant Argument might look very much like the argument Joyce (2005) gives, and I do not (here, at least) want to quarrel with anything Joyce has to say about the relevant motivational psychology. I do, however, want to stress that the Subjective Warrant Argument as I set it up is amenable, in a way Joyce's argument *might* not be, to a realist treatment of the target discourse (for more on which, see Section 5, below). I shall not

The Weak Argument (and the Weakness of the Strong Argument)

Even if you do not think that the prospects are good for defending the premise of the original (Strong) Subjective Warrant Argument according to which our moral beliefs are or would be *generally* subject to lack of subjective warrant, there is a weaker version of that argument which requires neither that premise nor the Safety Net Idea. The Weak Subjective Warrant Argument concludes not that we have reason to think it good if our acceptance of moral claims is *typically* nondoxastic and not doxastic, but just that we have reason to think nondoxastic moral commitment good.

The Weak Argument is in the spirit of van Fraassen's argument that empirical adequacy is a *permissible* standard of acceptance (van Fraassen, 1980). On the basis of the thought that our nondoxastic moral commitments would be sufficient to play action guiding roles in important high stakes situations in which morally guided action is extremely desirable, we have reason to say that a person who nondoxastically accepts some moral claim is, to at least that extent, doing well. Of course, they might accept a *bad* moral claim (see Section 5, below), but the worry that a person might have dangerous or otherwise undesirable moral commitments is just as much a worry with respect to a person's moral *beliefs*, of course. The point is just that to the extent that there is *some* (range of) moral commitment(s) which it is good for a person to have (and for them to have in such a way as to make it likely that they will act on that commitment (those commitments) in difficult circumstances, or in those circumstances in which a great deal rides on doing

discuss this any further here, but note that as Joyce is an error theorist it does not occur to him to argue, as I do here, that there can be a good in the nondoxastic acceptance of *true* claims, and certainly not in relation to morality.

the right thing), nondoxastic moral commitment is just as good as moral belief.

Of course moral beliefs will be preferable to nondoxastic moral commitments in contexts for which we are interested in a person's *epistemic* performance. That is why I have sought to defend normative fictionalism (or at least its nondoxasticist element) for our 'ordinary' or 'every-day' moral judgements and commitments only. Moral *philosophy*, where that is understood as related to every-day morality as the mathematics that goes on in university departments relates to every-day arithmetic, is plausibly such a context in which we ought to have and be interested in our moral beliefs, for when we *philosophise* about morality (and particularly when we do *metaethics*) we are interested in truth in a way that is often and quite reasonably eclipsed by our interest in the pragmatics (in the broadest sense, which includes the practical business of *being a particular sort of person*) of the morality of every-day life. Both the Strong and Weak Subjective Warrant Arguments leave room, though, for the fact that we had better not abandon belief as the appropriate mode of acceptance for genuinely truth-interested practices of enquiry. All they seek to do is to show that there is *something* good about nondoxastic acceptance of *some* sort of claims for the purposes of *some* practices in which our acceptance of those claims is important.

5. Realistic Fictionalism and the Subjective Warrant Argument(s)

I argued in Chapter Two that there is nothing pragmatically incoherent in both believing and nondoxastically accepting the same claim (even at or over the same time, and even when there is no doubt about whether the believed and nondoxastically accepted claim is the same one). The point for our purposes now is that we might be persuaded by (either the Strong or Weak) Subjective Warrant Argument on its own terms, but think that it is

undermined by the case for moral *realism*: if we are realists in our moral philosophy, must we not be bound to accept moral claims, even in every-day life, according to their *truth*, which is to say, must we not be bound to believe or not believe moral claims, but never to ‘merely’ nondoxastically accept them?

I do not see the force of that thought. We quite reasonably go about our lives in all sorts of ways which involve accepting things in a truth-*disinterested* way, even though we happen, in contexts where truth-interestedness *is* called for, to think that those things are true. Take, for example, a philosopher, Barney, who works on the metaphysics of fictional objects. Suppose Barney is a *realist* about fictional objects: he thinks that there exists (or at least that there is) an object which is Dorian Gray (though he is not committed to thinking that that object has the properties ascribed to Dorian Gray by Oscar Wilde; see Sainsbury, 2009 for a good survey of the various options for realism here). Barney likes to read fiction, and indeed that is one of the reasons he works on the metaphysics of fictional objects. But when Barney curls up with a copy of his current bedtime read, he accepts that, for example, Dorian Gray leaped to his feet, with flushed cheeks and burning eyes,²⁰ and thereby accepts that there *is* a Dorian Gray, *because that is required for engaging with the novel*; he would accept that there is a Dorian Gray in *that* context if he were *not* a realist in his metaphysics of fictional objects. In short, his acceptance of fictional claims, like all of our typical acceptance of fictional claims *qua* fictional, is not a matter of belief, for he does not accept the existential claim about Dorian Gray claim *because* he thinks that it satisfies the truth norm (or any truth-associated norm), even though *incidentally* he does believe that there is an object which is Dorian

²⁰ Wilde [1891]: p50

Gray. So there is no general reason to suppose that realism ‘crowds out’ nondoxastic acceptance.

That being the case, a moral realist might very well adopt a different attitude outside the confines of their study, or might carry their belief in a range of moral claims outside the study and nonetheless rely, for their day-to-day practical deliberation, on their *nondoxastic* acceptance of those claims. (Note that the ‘every-day’ nature of the deliberation and action I am concerned with here does not undermine the idea that some of the deliberations and actions in question are ‘high-stakes’: if we must make a literally life or death decision in the course of our lives outside of the study, we need not (in order to deliberate reasonably) retreat to the study to *philosophise* about it – there may be no time, or we may reasonably feel that that would not be appropriate anyway. In any case, life or death decisions (for example) are more common outside of the study than within it!)

Realistic fictionalists are, at any event, nicely placed to address a worry we might have about the Subjective Warrant Argument, or *any* argument for evaluative nondoxasticism. We might feel that nondoxastic acceptance – especially on the Character-Centred model I have proposed, with its relentless emphasis on ‘internal’ norms which specify *whether* someone accepts *P* and not whether they are *right* to – must threaten to push the fictionalist into an ‘anything goes’ attitude to our every-day moral commitments: if all that matters is whether some *pragmatic* good is served by accepting some moral claim (the good associated with our moral commitments achieving some influence over our *actions*), then there is no restriction on what, so long as it serves that good, a person accepts. And if the idea is just that it is good to have *some* moral commitments with sufficient motivational oomph and deliberative clout to play action guiding roles in high stakes situations, then so long as a person is strongly enough

committed in an action guiding way they are beyond criticism in respect of their having those commitments.

This worry is too quick anyway. Even a non-realistic fictionalist can plausibly repel it by noting that the pragmatic good in question is not just the good of morally guided action in high stakes situations *simpliciter*, but the good of acting *well* in some way, the likelihood of which is increased by acting on moral reasons of a restricted sort. The particular moral commitments it is good (and bad) to have, then, will depend on the role such commitments might play in achieving the goal of acting well by whatever measure of 'well' the fictionalist prefers.

Whilst any fictionalist is able to propose some candidates for what determines the sort of morally guided action which amounts to acting well, the realistic fictionalist has two very obvious and seemingly quite satisfying routes available to repel the 'anything goes' challenge. They might say that the actions which constitute acting well are those which achieve some genuinely moral good – realistic fictionalists, after all, need not scruple at positing states of affairs, actions, or anything else with full-blown moral properties. As fictionalists, they will say that our acceptance of particular moral claims need not be linked to any awareness of those moral facts which make particular actions morally good. But they will be able to insist, quite coherently, that as an *external* assessment of particular actions we are free to regard some actions as good, and those moral commitments (doxastic or nondoxastic) which are liable to promote those actions as good.

Or the realistic fictionalist might just settle for rejecting the 'anything goes' challenge by pointing out that since some moral claims are *true*, we can approve of those who are committed (even nondoxastically) to moral truths and deprecated the commitments of those who are committed to moral falsehoods. That would seem to be a reasonable and decisive end to the 'anything goes' challenge. But they might or might not then go on to say that

the availability of moral truths is relevant to the sorts of actions we have reason to want to promote, in that we have reason to want to promote actions done for the right moral reasons. Such a (partial) specification of the actions we can think of as acting well (and therefore as deserving our approval as pragmatic ends) seems to require the possibility of there *being* good moral reasons, so that specification of the sorts of action we have reason to approve is a response to the question about what acting well amounts to which is available to the realistic but not the non-realistic fictionalist.

6. Conclusion

Moral fictionalists of at least one type think that our acceptance of moral claims is or might be a nondoxastic attitude. A question arises, though, as to whether it is good that that is the case, or whether it would be good for that to be the case. Reflecting on the seeming facts about what moral *beliefs* are like, or would be like, in respect of their deliberative and action guiding roles in the highest-stakes – and often most morally significant – circumstances suggests that it *would* be good if our moral commitments were to be nondoxastic, for nondoxastic commitments are typically not subject to the undermining factors which threaten the deliberative and action guiding efficacy of our beliefs. That is not to say that if our moral commitments are currently typically doxastic we have decisive reason to change, nor that if our moral commitments are currently nondoxastic we have no decisive reason to change. But it is to say that there is a serious debate to be had about what the relative goods of doxastic and nondoxastic moral commitment are, and that fictionalism is not by any means obviously a fall-back or second-best option.

In the previous section of this chapter, I introduced the idea that the realistic fictionalist – but not the non-realist fictionalist – might usefully appeal to the thought that there are some moral commitments which are the right commitments to have in virtue of the moral facts, in response to the worry that, granted the Subject Warrant Argument establishes that it would be good for *whatever* moral commitments we have to be nondoxastic, it does not matter *which* moral commitments we ought to have for all the fictionalist has to say. In the next chapter, our topic will be related: What is there to say by way of vindicating the intuitive and appealing thought that there *are* some moral commitments we ought to have at all? It will turn out, again, that the realistic fictionalist is better placed than the non-realist fictionalist to satisfy the demand that we say something about this.

Vindicating Realism

1. Introduction

In the previous chapter, I argued that it would be good in at least one way if our moral commitments were to be as I suggested, in Chapter Four, they might be and often are. And I have insisted, throughout the preceding chapters, on a sharp distinction between our *ordinary* practice and the sort of thinking *about* that practice and the discourse it involves which is characteristic of the sort of philosophical reflection which leads us to formulate, defend and argue over theses concerning the semantics, metaphysics and epistemology of our ordinary practices. I want, in this chapter, to focus on that distinction in more detail, and to show more explicitly (what has already been hinted at the end of the previous chapter) how, for the case of our ordinary moral commitment and deliberation with which the second half of this thesis has been concerned, *realistic* fictionalism might be an advance upon the usual forms of *non*-realist fictionalism.

I shall be interested in the prospects for vindicating the thought that there are some moral commitments we ought to have.¹ As short-hand, I shall usually refer to this as the prospects for vindicating our moral commitments, or for vindicating morality. Extremely few people do not think that there are some moral commitments we ought to have, and it is one of the tasks facing

¹ We may read this phrase ('There are some moral commitments we ought to have') in either of its two available ways: as implying that we know which the particular commitments are which we want to vindicate; or as saying just that there *are* some commitments we want to vindicate, but implying no more than that we know that we think it is a good thing to have some moral commitments rather than none, but are as yet unsure of which are the particular ones for which vindication is desirable. In the discussion which follows, I shall certainly not be taking a stance on which those commitments which we ought (for some reason) to have are.

moral philosophy to explain what, if anything, is right about that thought. I will argue that the most satisfying way of vindicating the thought that there are some moral commitments we ought to have employs a *realistic* fictionalism of the sort I have been suggesting.

I am not going to be presenting any argument for moral realism here, though. My interest in this chapter is with the vindicatory adequacy of various sorts of metaethical theories, with how adequately various sorts of metaethical theories do justice to our sense that there are some moral commitments we ought to have, a sense which is not *necessarily* tied to our sensitivity to the truth of any form of moral realism. Indeed, as we shall see, several metaethical theories seek to vindicate our moral commitments without recourse to any distinctively moral facts or truths. A major theme running through this thesis has been that if you are already fond of a particular argument for moral realism (or some other sort of realism, for other sorts of fictionalism), you can avail yourself of the important insights fictionalism has to offer about the nondoxastic character of our acceptance of some claims as it often is and perhaps ought to be. If you are attached to some argument for moral realism – or if you insist that there just must be a good argument for moral realism, because it must be true – then you will already, if you accept the suggestions made earlier in this thesis, be disposed towards realistic fictionalism, and that is all well and good as far as I am concerned (it is more or less my own position). This chapter is not intended to preach to the converted on that score, nor to convince anyone not yet convinced of the truth of moral realism that moral realism is true. It is intended to persuade those whose commitment is just to a metaphysically agnostic form of fictionalism such as the one I have developed to account for how our moral commitments often are and generally ought to be that if we are interested in vindicating the common thought that there are some moral

commitments we ought to have, we have reason to embrace realistic rather than non-realist fictionalism.

Just as everything is notoriously, and vacuously, similar to everything else in *some* respect, everything is, equally vacuously, good in *some* respect. Vindicating something is not just a matter of finding some respect in which it is good; it is finding some *relevant* respect in which it is good. Relevant to what, though?

We want to vindicate our moral commitments *as the sorts of things which play the roles for which we value them, and as the sorts of things we value them being*. We have no interest in vindicating our moral commitments if it turns out that the basis on which we vindicate them does not also vindicate our treating them as valuable. So the value we attach to a person's having some moral commitments is the source of one important constraint on what is to count as an adequate vindication.

In what follows, I shall move between talk of the moral commitments we ought to have and the goods in virtue of which, on various putatively vindicatory accounts, we ought to have them. A properly careful treatment of these issues would, of course, need to account for the fact that what is good does not straightaway entail what we ought to do: some good might be insufficiently good, or insufficiently significant, to mean that we ought to do what is required to promote that good, all things considered. But in what follows, it is to be understood that when the relationship between a good and the commitments we ought to have is invoked, the idea is that that good is sufficiently good and important that it means, all things considered, that we ought to have the moral commitments which would promote that good.

It might seem that if the vindicatory accounts I am to be concerned with are to be understood as invoking some good which is sufficiently good and important that it means, all things considered, that we ought to have the moral commitments which would promote that good, then the only way to

seriously criticise those accounts would be to doubt that that good *is* sufficiently good and important.

I will argue that what I'll call *instrumental* vindications of our moral commitments are bound to be inadequate. But I do not doubt – to give a glimpse of what is to come – that the regulation of our behaviour in particular ways *is* sufficiently good and important. So perhaps I must concede that whilst the instrumental vindications I intend to criticise do not work in quite the way we might expect given our typical attitudes to the value of morality, they do nonetheless invoke the right sort of value to vindicate our moral commitments simply *because* they invokes a good which is sufficiently important to mean that, all things considered, we ought to have some moral commitments.

Vindications, though, can be more or less satisfying, and a vindication which ties the all things considered reasons for having some moral commitments to more of the things we value is, thereby, a more satisfying vindication. I shall argue that a realistic version of the Character-Centred fictionalist theory can vindicate our moral commitments by invoking non-instrumental goods which *also* mean that, all things considered, we ought to have some moral commitments. There is, then, a stand-off between what we might choose to call adequate vindicatory accounts. But, I shall argue, the realistic Character-Centred theory can give us everything the instrumental theories can give us, in addition to those things we might want of a vindicatory account which those instrumental theories cannot give us. So it seems reasonable to say that the instrumental theories are not *all things considered adequate*. That, then, is what my talk of 'adequacy' throughout this chapter should be taken to mean.²

² I am sensitive to the worry that the *adequacy* of a theory is not really relative to the availability of better theories: adequacy is, of course, relative to the facts to the available evidence, but surely if a theory is adequate relative to the facts and the evidence then it

2. Error Theory and the Challenge to Ordinary Commitments

Having, he thought, shown in the first chapter of his book, *Ethics: Inventing Right and Wrong*, that there must be some very queer sorts of facts in the world for our ordinary moral claims to be true and that there are no such facts (at least for all our sensibly informed metaphysics and epistemology can say), Mackie faces an interesting problem, with which any error theorist is bound to be faced: must we say that our ordinary commitments are to be abandoned, or that there is something to be said by way of vindicating them which doesn't rely on their being true?

Mackie says that:

[W]e have to decide what moral views to adopt, what moral stands to take. No doubt the conclusions we reach will reflect and reveal our sense of justice, our moral consciousness – that is, our moral consciousness as it is at the end of the discussion, not necessarily as it was at the beginning. But that is not the object of the exercise: the object is rather to decide what to do, what to support and what to condemn, what principles of conduct to accept and foster as guiding or controlling our own choices and perhaps those of other people as well. (Mackie, 1977: p106)

This ought, I think, to be a rather inspiring passage for an error theorist, for it expresses an intention to take first-order morality seriously despite metaphysical and epistemological commitments which threaten the literal truth of that morality, and thus resists moral *nihilism*, which any reasonable person would deprecate, and which is all too often assumed to simply follow in the wake of metaethical scepticism. No matter, from this point of view,

remains *adequate* even if there is a better available theory – it still does the job for which it is intended as well (or badly) as it would do if no better theory (indeed, no other theory at all) were available, after all. So I am willing to concede that the word 'adequacy' might be infelicitous here. If it does seem unacceptably infelicitous, the reader is free to substitute some terminology of their own to stand for that property of a vindicatory account which consists in its satisfying our interests in having a vindication at all.

that we have to *decide* what moral views to adopt and what moral stands to take rather than *discovering* them; the role of those moral views and stands in regulating our behaviour towards one another is just as significant, on Mackie's view, as if they were discovered in the fabric of the world or the structure of practical reason itself.³

But having abandoned the idea of discovery in favour of decision, Mackie's project of building first-order morality back up from the damage his error theory seemed to do it raises questions which, in Mackie's hands, are not adequately addressed. The project is one of vindicating (some of) our first-order moral commitments from the point of view of one for whom, in answer to the question '*Which* claims (or principles) should I accept?', the simple answer 'The true ones, because they are true' is not available. What, then, *is* Mackie's answer?

³ In a way which invites further consideration elsewhere, this emphasis on deciding rather than discovering is apt to remind us of some of the things neo-pragmatists (avowed disciples of Dewey, in particular) such as Rorty (e.g. Rorty, [1985]) and the later Putnam have said about values in general and moral and political values in particular. But we ought to be careful before assuming that Mackie's point here is congenial to anything characteristically neo-pragmatist. Certainly on Putnam's version of neo-pragmatism, the denial of the fact/value dichotomy is, if we are permitted to retain the vocabulary of 'facts' and 'values' for a moment, as much about showing that values are facts in any respectable sense of 'facts' as that facts are value laden. Indeed, Putnam seems to reject Mackie's arguments from queerness partly for this reason (Putnam, 1994: pp156-9). And Putnam is keen to point out that on his understanding of the pragmatists' (in particular, Dewey's) insights, it is a mistake of Rorty to disown talk of 'getting the facts right' (see, e.g., Putnam, 2002: ch6). For what it's worth, I incline towards Putnam's understanding of what the classical pragmatists were up to rather than Rorty's, and I think Putnam's neo-pragmatism is far more attractive than Rorty's; but for our purposes here, it suffices to note that accepting that values are decided upon rather than discovered falls some way short of committing us to anything as contentious as Rorty's 'ironism' (see Rorty, 1989).

It might have seemed to some readers that the theory of ordinary moral commitment developed in Chapter Four amounts to something like Rortian ironism. I have no desire to enter into exegetical debates about how close the Character-Centred theory comes to what Rorty actually thought, but it certainly is not the case that the *grounds* for the Character-Centred theory which I proposed are anything like the neo-pragmatist grounds upon which Rorty's theory of value is based. And, if I am right about the desirability of *realistic* fictionalism, the most satisfying finished version of the Character-Centred theory is robustly realist in a way which would, I take it, have appalled Rorty.

His answer adverts to the job a first-order morality can do for us: in light of our limited sympathies, our joint need for cooperation, and the scarcity of the needs and comforts of life, we must regulate our actions which affect others, for the good of all. Nothing in that idea yet determines what the *nature* of our moral commitments (*qua* mental states), by means of which such regulation is to be achieved, ought to be, and Mackie does not tell us. So perhaps his thought is that there is no special problem here, and that we can, once armed with *some* vindication for (some of) our moral views, have done all that is required in so far as the vindication of (some of) our moral *beliefs* is concerned.

But is that really what a serious error theorist ought to be content with? An error theory entails that, for any of its target claims *P*, *P* is false. On the most intuitive and reasonable logical assumptions that entails, of course, that not-*P* is true. Now, the error theorist need not subscribe to an epistemological principle as implausibly strong as to say that a person ought to believe every entailment of a true theory. But they might very well subscribe to the far more reasonable principle that a person ought not to believe anything which contradicts an entailment of a true theory. And if they do, they must agree that a person ought not to believe *P*, given that their theory entails not-*P*. Such an error theorist, then, would surely be embarrassed to recommend, on *instrumental* grounds, that we *believe* some moral claims.⁴ Of course, the *epistemic* deontology I am tacitly invoking here is not strictly incompatible with the view that there are instrumental reasons for beliefs, nor even with the view that there might be *overriding* instrumental reasons for beliefs. Perhaps it is, all things considered, a good

⁴ 'Instrumental' and 'non-instrumental' are, like so many terms employed in this thesis, contested in respect of their meaning, and are employed in different debates with sometimes quite subtly different meanings. Here, I mean by an 'instrumental' vindication a vindication which depends upon some instrumental good, and by an 'instrumental' good a good which is good in virtue of its promoting some other good.

thing to believe even contradictions (never mind just propositions which contradict entailments of true theories). I am not going to take a stand on whether it is. In any case, we are *forced* to admit that we ought to believe moral claims which are strictly false on instrumental grounds – something which on any reasonable view counts as *some* cost – only if there are overriding reasons to accept those moral claims *and there is no other way of accepting them adequate to those reasons for accepting them*, by which I mean that, whatever instrumental grounds there might be for our accepting some false moral claims, those will only be grounds for our needing to *believe* them if they are not also (just as good) grounds for *nondoxastically* accepting those false moral claims. (Compare Joyce, 2006: ch5)

This does not seem to be an issue to which Mackie is sensitive, but of course it is an issue to which fictionalists are characteristically responding: on the currently most usual fictionalist model, an error theory (or at least an argument for agnosticism, which similarly undermines our licence to believe with confidence) is pitted against the instrumental good of accepting the erroneous claims and, by way of compromise, it is suggested that the (epistemically and pragmatically) responsible person will, in the course of the practice in question, accept but not believe the target claims. This does not incur the cost (whether or not that cost is decisive) of endorsing false beliefs, for it is not beliefs which are being endorsed, precisely because they are (or would be) false. (See Joyce, 2001; 2005 for a paradigm example of this fictionalist strategy.)

But the *instrumental* vindication of our moral commitments, whether the vindicated commitments are beliefs or something nondoxastic, is subject to what I'll call a 'Wrong Sort of Value Objection'. According to that objection, it is not enough that there be *something* to say about the good of accepting some claim; we have *reason* to require a vindication of our moral commitments, which reason is not met by an instrumental vindication. The

objection is that instrumental reasons for accepting some moral claims are not (and do not relate in the right way to) the reasons we have for seeking to vindicating them, so the vindicatory project which we have reason to pursue does not succeed, or at least does not succeed in the right sort of way, for it fails to deliver what was required of the project (fails, that is, to meet its *raison d'être*).⁵

3. How a Vindication Can Be the Wrong Sort of Vindication

Error theories have the power to be rather troubling, and when they are it is because they are about things we value in some way. An error theory about witches, for example, is not likely to be disquieting to many of us in philosophy departments of universities in industrialised western societies, but in times or places where belief in witchcraft is part of a conception of what the world is really like (its forces, and the parts we can play in taming and being subject to them), such an error theory represents a serious challenge: the beliefs upon which those people's understanding of what the world is like (a basic and important sort of understanding) are partly based are, according to that error theory, false and, if they accept that theory, irrational, so whatever reason there is to think that an error theory about witches is correct is *ipso facto* a threat to a set of commitments which are not easily given up, or which are at least not given up without (in the absence of a replacement conception of what the world is like) losing something which we value.⁶

⁵ It is independently plausible that we generally need to understand the *point* of a theory; isn't that what we try to convey in teaching theories, and try to grasp in learning them?

⁶ There is, of course, an interesting set of issues (in which philosophy and anthropology might be thought to intersect) about the extent to which a classical error theory, with its fully representational semantics, really does undermine whatever value a set of commitments has for a particular group of people: if it turns out that witch-talk is not supposed to be indicative of the literal beliefs about the world of those who talk of witches, but rather that

Witchcraft error theories are not, as I said, very troubling for most of us. We probably accept such an error theory, and it probably *reinforces* our conception of what the world is like rather than undermining it. But there are plenty of error theories about things we do find it very tempting to believe in, error theories about ordinary objects like tables and chairs, for example.⁷ Such error theories are easily dismissed by the man on the Clapham omnibus, and by a good many professional philosophers, but the point is precisely that they are easily dismissed *because* they would, if accepted, undermine something (our conception of what sorts of things the world contains) which we are typically unwilling to have undermined.

We are also typically unwilling to have our moral commitments systematically undermined, and with good reason. Perhaps what we have good reason to value and to resist the undermining of are our moral beliefs, and the reason we have to value them is that, like our beliefs about ordinary objects and the impossibility of witches, those beliefs are part of our conception of what the world is like: the world contains things such as tables and chairs, could not possibly contain witches, and is such that some things are wrong or good. But we do not need to think of the moral commitments we have reason to value as beliefs about what the world is like in order to understand what we value in them when we seek to vindicate them in the face of an error theory.

the witch-talk is part of a game the purpose, and value, of which is to do with a person's willingness to express their solidarity with their fellows' way of talking about the world (perhaps a way descended from a time when people *did* believe in witches, but a way which is now to do with the value of tradition and not the value of being strictly accurate about the way the world is in day-to-day talk), then the error theory will not, in fact, undermine anything of the value that witch-talk possesses for that group. But then their commitments to witches will not be most charitably interpreted as beliefs anyway (at least on the model of belief and acceptance with which we are working in this thesis).

⁷ See Thomasson, 2007 for a defence of our commonsense notion of ordinary objects against various sorts of error theories.

Our moral commitments do far more for us than just limit a domain of our conception of what the world is like, even if they do that, and their value is not, as I have suggested in previous chapters, tied to *that* function of theirs – we do not value our moral commitments (our own and others’) because they tell us what we think the world is like (even if they *do* tell us that), but because they tell us what *we* are like.

Our beliefs about what the world is like in non-moral ways can tell us about what we are like, too, of course. If we learn, of a person who is in a position to know that the Earth is roughly spherical, that in fact they believe the Earth to be flat, then we might very well not just take ourselves to have learned something about what they think the world is like; we are likely to conclude with some justification that they are the sort of person who either doesn’t take getting things right about the natural world seriously, or who is seriously lacking in some epistemic virtue(s), perhaps that they are guilty of some form of inattentiveness, carelessness or other sort of negligence. But that does not mean that the *value* of our beliefs about the basic physical properties of the Earth is to do with the role those beliefs can play in our judgements about what people (ourselves and others) are like.

Our moral commitments, on the other hand, play an important role for us in our judgments of character, a role which, when seeking to vindicate those commitments, we must defend. It is largely a matter of indifference whether we think of someone as particularly competent in respect of knowing what the physical properties of the world are, even though it is important to us that there be a vindication of those commitments which are central to our conception of the world. But it is important to us whether we think of someone as having the right sort of moral commitments. Our attitudes to them in respect of whether we want, for example, to be friends with them, whether we are willing to help them in some endeavour of theirs, whether we shall recommend them to a lonely friend as good partner

material, whether we forgive some perceived misdemeanour, whether we trust them, and all manner of other things are (rightly or wrongly) typically in large part a function of our attitudes to a person's moral commitments, to say nothing of all the attitudes to *ourselves* we adopt on the basis of our views of our own moral commitments. These are sources of value for us of our moral commitments. And if a vindication of our moral commitments is to succeed (if it is to defend the value of those commitments in the right sort of way, to sanction our confidence in them having *that sort of value*), it must not just give us *some* reason to go on with our moral commitments – it must explain that those commitments have that sort of value for us.

Let us return then to the vindications offered by the standard error theory accepting type of fictionalist, the fictionalist who, like Field or Joyce, is impressed by the arguments for treating their target discourse fully representationally but is equally impressed by some argument to the effect that our beliefs about mathematics or morality are bound to be systematically false, or at least unjustified. Let's take Field's mathematical fictionalism first. Field thinks that our mathematical practice is vindicated by the fact that it is a relatively convenient way of deriving the results (and only the results) of a literally more accurate nominalistic physics, a calculation and proof method which is less unwieldy and more easily applied than the more metaphysically respectable physical calculus he develops to replace traditional mathematical physics (Field, 1980; 1989). In a similar vein, Yablo, 2005 thinks that our mathematical practice is vindicated by the power of mathematics not just to calculate and prove, but to *represent* things – not to represent ontologically contentious mathematical things, but to represent less contentious things in uniquely useful *ways*. What both these versions of mathematical fictionalism have in common is their confidence in the propriety of our mathematical practice on the basis of its usefulness.

Such a vindication of our mathematical practice in the face of an error theory which, from the fictionalist's point of view, seems compelling is arguably perfectly adequate (so long, at least, as the trumpeted advantages in utility really are afforded by our mathematical practice). After all, we might never have really ever seen any great value in mathematics other than as a way of calculating and proving (and perhaps representing) things about the world anyway, even if we accepted that in fact there are mathematical objects and our mathematical claims are true: it is not an unbearable stretch of the imagination to think that the only thing we have typically *valued* about mathematics is its enormous power to (help us) predict and build things. And if that is all we have typically valued mathematics for, the fictionalists' utility-invoking vindication seems entirely germane. Nothing, at least nothing in the way of explaining what reason we have for engaging in mathematical practice, is lost from what we had before.

The fictionalist vindication of our mathematical practice, then, is successful (if it is successful at all) because it preserves what is of value in that practice, or at least what we typically value in it. True, some *philosophers* who are mathematical realists might feel short-changed by the fictionalist's utility account of the value of mathematics, insisting that there is intrinsic value in the truth, and that the *usefulness* of mathematics is no substitute value for that intrinsic value to which the error theorist has no right to appeal. For such a philosopher, no *accommodation* of the error theory will suffice for vindicating our mathematical practice and our mathematical commitments in the right sort of way; we must *reject* the error theory, and re-entitle ourselves to the idea that the value of our mathematical commitments is explained by their truth. Never mind whether anyone, philosopher of mathematics or otherwise, would be right to say that there is intrinsic value in the truth (let's assume for the sake of argument that they would be); what is at stake here is whether it would be right to say that the value of our

mathematical practice and commitments depends upon that value. Our supposed mathematical realist would presumably be entitled to say such a thing, since for them there *are* mathematical truths the intrinsic value of which, *qua* truths, requires something of our mathematical practice and commitments (namely that we believe the truths, and that our practice afford a privileged place to those true beliefs, because they are intrinsically valuable *qua* true beliefs). On such a view, which grants the intrinsic value of true belief, it follows from there being mathematical truths that no vindication of our mathematics which does not mention such truths (and indeed denies them) could capture the full range of reasons we have for valuing mathematics. The utility vindication, then, is, on such a view, doomed to fail as an *adequate* vindication; it will be too revisionary of such a philosopher's ideas about what the value of mathematics is.

But the mathematical realist here is playing the game the wrong way around: having not been convinced at all by the error theory in the first place, and still believing, therefore, in mathematical truths, they are pointing out that accommodating the error theory by vindicating our mathematical practice by some means other than adverting to the value of the truth is inadequate. The interesting question, though, is how adequately the fictionalist's vindication of our mathematical practice satisfies our less theory-laden ideas about the value of mathematics. If the seemingly most reasonable idea about the value mathematics has for us makes no mention of the intrinsic value of the truth, and the fictionalist's vindication leaves the story we want to tell about the value of mathematics intact, then we have reason to think of the fictionalist's vindication as adequate *qua* vindication of our mathematical practice and the reasons we have for engaging in and valuing it.⁸

⁸ If we think that there is an intrinsic value in the truth, or in having true beliefs, then we can vacuously append that value to the account we give of the value of *any* sort of commitment

4. The Wrong Sort of Vindication of Morality

If the moral fictionalist can tell a story about the propriety of our moral commitments in which those commitments have the same sort of value as we take them to have independently of any concern with the metaphysics of morality, we can fairly be said to be in possession of a successful vindication, a vindication of our moral commitments which does not require us to be extremely revisionary in our conception of what the role and value of those commitments are. But, though it is plausible that the mathematical fictionalist is entitled to give an error-theory-accommodating account of our mathematical practice which, in accommodating the error theory, describes the value of that practice in terms of its utility (because it is plausible that utility is what we value in our mathematical practices, anyway), it is not so plausible that the moral fictionalist can give a satisfying account of our moral commitments which, accommodating an error theory, describes their value in terms of anything that does not ultimately rely upon the truth of those commitments (or of some of them, at least) or some other closely related distinctively moral truths.

The objection to putative vindications of our moral commitments which advert 'all the way down' to utility or some other instrumental value

which we happen to think of as commitment to the truth. But the vacuity of such an appendage reveals that the interesting debate is elsewhere: we often want to know whether, in order to give an adequate account of some domain of commitment, we are required to treat that domain realistically (whether, e.g., religious commitments are vindicated only as beliefs in an actually existing deity, or whether those commitments are vindicated even in the absence of such a deity, or in the face of necessary ignorance of Him), precisely because we are more sure of what to say about the practices involving the target discourse than we are of what to say about the metaphysics of that discourse.

That, of course, is a purely methodological point. But aside from that point, I think it is clear that the assumption which we have been granting for the sake of argument here – that there *is* an intrinsic value to the truth and/or to believing the truth – is far from obvious and that any argument for it is bound to be less uncontentious than an argument that, for a *specific* domain, we value commitment to the truth for some reason.

is *not* that what I'll call an error-fictionalist vindication can say nothing plausible about what is or would be *good* about our moral commitments. (Indeed, the Subjective Warrant Argument I gave in the previous chapter is an argument to the conclusion that some nondoxastic moral commitments *are* good to have, and that argument did not essentially involve any premises or presuppositions that the error-theorist would be bound to reject.) The point is that a vindication which depends upon just *something* good about the vindicated phenomena is not automatically satisfactory with respect to the interests we have in vindicating that phenomena in the first place.

A vindication which rests on some good which fails to match up with anything we value about vindicating that phenomena at all is bound to seem like an *irrelevance*. It would, perhaps, be irrelevant to advert to the truth of our mathematical claims if we are interested in vindicating our mathematical commitments, if what I said about the plausibility of their having just utility-value for us is right. If what we value about our mathematical commitments and practices is their utility, what we require of a vindication of those commitments is some reassurance (of the type supposedly secured by Field's (1980) safety result, perhaps) that they are as useful as we hope they are. Notwithstanding that we might happen to take a stance on the metaphysics of mathematics, is it not necessarily *false* to say that *another* good thing about our accepting some mathematical claims is that they are true, and that it is, *ceteris paribus*, good to believe the truth about things we are interested in. Nonetheless, it is largely *irrelevant* to the required vindication, the rubric of which mentioned only the requirement that some reason be given for thinking that mathematics is as useful as we hoped it would be (hoped it would be, because that is what we really value about it).

Mackie seems, implicitly, to be sensitive to the demand for relevance when he opens the second section of his book (Mackie, 1977: ch5) with a discussion of the role of morality. Surveying the views of Hobbes, Hume and

Warnock (amongst others), Mackie finds what he takes to be a compelling conception of morality as instrumentally useful, a means by which we can better live together. This, Mackie evidently thinks, is in large part what is valuable about morality *regardless of whether or not its claims are, as his error theory maintains, false* – the point seems to be that it doesn't matter that all the interesting moral claims are false, for their value never depended upon their truth in the first place. So the first-order morality developed throughout the remainder of Part Two of *Ethics: Inventing Right and Wrong* is vindicated by its ability to play the instrumental role in the regulation of our actions and lives for which we would, on this instrumentalist view, value any morality anyway.

Notice that had Mackie not taken any trouble to motivate the idea that a morality is valuable because of its instrumental role, pointing out the instrumental advantages of any particular morality would be something of a *non sequitor*, since it would remain an open question whether *those* advantages have anything much to do with the reasons we have for being interested in morality in the first place – much as it would be something of a *non sequitor* to advert to the fact that accepting some particular moral claims makes one more attractive to the opposite sex: that might be a *good*, but that does not mean that it has anything to do with what we are interested in when thinking about the propriety of our moral commitments. It is, I think, to Mackie's credit that he proceeds in this way, recognising as this seems to imply that vindications require more than just the existence of some good or other.

But if Mackie has the right methodology for a vindication, he does not seem to have the right idea about what an adequate vindication of our moral commitments specifically requires. For Mackie, and for his fictionalist successors, the vindication of our moral commitments (be they beliefs or, as they are for the fictionalists, nondoxastic modes of acceptance), requires just

that there be some plausible account of how those commitments play the action guiding and regulating roles required of them.⁹ Such roles are important, and they are *part* of what we value about morality. But they are not all we value about morality, and they are not seemingly what we value most centrally about it. The remainder of what we value about our moral commitments – the remainder which Mackie and his instrumentalist sympathisers leave out, but which is at the heart of what we most value about morality – is the importance of a person’s having particular sorts of moral commitments regardless of their actions.

To see that the instrumental action guiding and regulating roles of our moral commitments are not the central source of the sort of value we typically ascribe to morality, consider a possible world in which the sorts of control over our actions typically affected, it is supposed, in our world by a person’s moral commitments is affected instead by a neurological chip remotely controlled by a central regulatory authority, since we are never in any of the sorts of psychological states which, in this world, amount to our having moral commitments. This ‘control chip’ need not fully determine our actions to any greater extent than our moral commitments are supposed to control our actions – perhaps the chip functions just to make us precisely as uncertain of the all-things-considered desirability of an action as our moral commitments would make us, or just as inclined to adopt a pro-attitude to some action as our moral commitments would – so there is no question of

⁹ It is perhaps worth noting here that the class of vindicatory theories I am thinking of as ‘instrumental’ is broader than just the class of theories which advert to morality’s *social* role. Indeed, I count Richard Joyce as an instrumentalist in the relevant sense, since he argues (in Joyce, 2001: chs.7 & 8; 2005) that moral commitments are *motivationally* important for us, which I take to mean that he thinks it a reason to have moral commitments that there are things we ought for our *own* good to get around to doing, as well as for the good of us all. Joyce is explicit in his rejection of most versions of the vindication of morality by appeal to the ‘pro-social’ evolutionary advantages of moral commitments, though (Joyce, 2005: ch5), and it is easy to see that, fashionable as such an approach currently is, the idea that morality is *evolutionarily* advantageous for our social coordination is not by any means the only (or, seemingly, the most plausible) way of appealing to the instrumental value of morality.

the envisaged case involving a person's reduction to a mere automaton. In fact, it is not far-fetched to say, having recognised the proscribed role envisaged for the control chip, that, if the chip turns out to be the best or the only way of affecting the desired regulation of behaviour that is supposed to be the source of value for our moral commitments in this world, the chip and all its works are a good thing, in that respect at least (for the possible world in which it has that privileged status; it would not follow it would be a good thing for *us*).

What is the difference between the control chip in its possible world and moral commitment in ours? The difference is not in the instrumental roles of these mechanisms: each serves to regulate the actions of all in the interests of each. Nor is the difference in the success with which these mechanisms achieve their instrumental goals: *ex hypothesi* each does just as well. Is there, then, no difference in the value we ascribe to each?

It seems clear that there *is* a difference in the value we ascribe to our moral commitments in our world and the value we ascribe to the control chip in its: ¹⁰ were we in the control chip world, we would not be inclined to see ourselves and others as praise- or blame-worthy in the sort of way we are inclined to see ourselves and others as praise- or blame-worthy on the basis of our moral commitments; we would not be inclined to see a person's tendency to behave in certain ways as linked in a particularly interesting way to their character at all, in fact; so we would not be in a position to do a whole raft of things which we greatly value, things which depend upon our being able to sincerely say of a person (ourselves or someone else) that they are doing, or refraining from, or have a disposition towards some sorts of

¹⁰ It will probably be obvious that this thought experiment owes something to Nozick's (1974: pp42-3) 'experience machine' thought experiment, which similarly employs the conceit of a mental state-affecting piece of sci-fi equipment to undermine a theory (utilitarianism, for Nozick) about the value of things, based on our intuitive (though still reflective) judgments about the relative value of particular types of sci-fi equipment-induced phenomena.

actions *for the right reason* (and *mutatis mutandis* that they have the wrong reasons for doing things). The psychological effects of the control chip are just too course-grained: the chip just statistically regulates our pro-attitudes and judgements of all-things-considered desirability, but it does not produce any distinctively *moral* pro-attitudes, and does not induce any particular considerations; a judgement of all-things-considered desirability does not necessarily involve a consideration of all things.¹¹

Such a neuropsychological mechanism could be of value to us only to the extent that our interest is in regulating our behaviour in a non-coercive way.¹² It could not be of any value to us with respect to its role in grounding judgements of praise or blame, or its role in grounding our judgements about whether a person acted for the right reasons (unless we have an implausibly thin notion of what acting for the right reason is, on which we act for the right reasons just in so far as we act on our pro-attitudes or judgements of all-things-considered desirability), for it plays no such roles.

If morality is just like the control chip in respect of what we value about it, then there ought to be no difference in the ways or extent to which we value our moral commitments and the ways or extent to which we value the control chip. But we are not inclined to value the control chip in the same

¹¹ It is crucial to recognise that what the control chip is supposed to do is to induce (in a statistically regulated way) judgements *of* all-things-considered desirability. Were the chip to induce reflection on or sensitivity to all the relevant considerations, and thereby induce an *all things considered judgement* of desirability, our intuitions might be very different from the ones I am pumping here. The difference between an *all things considered judgement* of desirability and a judgement *of* all things considered desirability is significant, though, for whilst the former necessarily involves consideration of all the relevant considerations, the latter might itself be an entirely ungrounded judgement. It is like (indeed, on some conceptions of reasonableness it is just a species of) the difference between a reasonable judgement and a judgment of reasonableness.

¹² Recall that the chip does not *compel* us to act or refrain from acting to any greater extent than our moral commitments would. A person's actions, in the control chip world, are just as much a function of their own pro-attitudes and judgements of all-things-considered desirability as our actions are; all that is different is that what explains their having those pro-attitudes and judgements of all-things-considered desirability is the control chip and its operation by a remote source, rather than (as in our own case, often) our acceptance of certain moral claims or principles.

way as or to the same extent that we value morality, so what morality and the control chip share – their instrumental role in regulating our behaviour to some extent – cannot be all there is to the value of our moral commitments for us.

5. An Adequate Vindication of Morality

The point of the control chip thought experiment is to pump the intuition that adverting to the instrumental role of our moral commitments cannot adequately vindicate them: what we are interested in respecting in a proper vindication of morality is not just our moral commitments' instrumental role in regulating our behaviour. What, then, *is* required of an adequate vindication?

What is required is an account of the propriety of our moral commitments in terms of the *non*-instrumental goods which moral commitment seems to involve. We value a person's morality not just for the role it plays in regulating their behaviour but because it is good to be a person who is committed to particular sorts of moral claims regardless of those commitments' role in action-guiding. It is *not* all the same to us whether a person helps old ladies across the road because they want to attract the admiration of others or because they think it is the right thing to do (or because they are properly sensitive to some features of the situation which are distinctively moral), and the way in which it is not all the same to us is not at all adequately captured merely by the thought that moral commitment or sensitivity is a more *reliable* mechanism of action regulation, the reliability of which deserves special respect. This is one of Kant's important insights (Kant [1785]: Sec.1).

The Character-Centred model I explained in Chapter Four can vindicate (or rather play a role in a vindication of) our moral commitments

in respect of their instrumental role in regulating our behaviour just as well as any other account of our moral psychology, for it is open to the proponent of the Character-Centred model, as it is open to the proponent of any model, to argue that there are some moral commitments the *content* of which is such as to be good in respect of that regulatory role. In general, plausible accounts of the *nature* of the mental states which our moral commitments consist in will be neutral with respect to whether or not there is any good arising from their action-guiding and regulating roles: in order to be plausible accounts of the nature of our *moral* commitments, such accounts must allow that (and go some way to explaining how) those commitments *do* play action guiding and regulating roles; but whether there are any commitments which we ought to have in respect of regulating our behaviour will subsequently depend upon whether the content of any of our moral commitments is or could be such as to affect a regulation of our behaviour which is itself of some value. Indeed this is, I think, why Mackie concerns himself (in the second section of his book) with working out (or making progress towards working out) what the *content* of a first-order morality which would adequately play the regulatory role of morality (a role the value of which is tied to our interests in living reasonably peaceful and cooperative lives) would be, without concerning himself with what the character of our commitments *qua* mental states ought to be.

So a proponent of the Character-Centred model as a model of what our moral commitments sometimes are and often ought to be is entitled to whatever vindicatory power is provided by a conception of the value of morality as instrumentally good in respect of its role in the regulation of behaviour: whichever particular commitments are supposed, on that conception, to be instrumentally good in respect of their helping us to get along with each other, or in respect of our managing to get around to doing what is good for us, are just as well vindicated by the Character-Centred

account of the *nature* of those commitments as by any other plausible account.

But as we have seen, more is required of an adequate vindication of morality than just to reassure ourselves of that instrumental good. In order to have defended the propriety of our moral commitments in terms of the right sort of value, we must have found a way of entitling ourselves to the idea that having some (sorts of) moral commitments is good in other ways, too. We must find grounds for saying that a person who does the instrumentally good thing *for the right reason* is better in some relevant respect (or that his action is better, or at least more praiseworthy) than the person who does the instrumentally good thing for some other reason.

Crucially, agreeing that this is required of an adequate vindication of our moral commitments, an account of their propriety in terms of the right sort of value, does *not* depend upon taking any particular stance towards consequentialism. The point I am currently pressing is a point about the value of our moral commitments, and not about any constraint on the content of morality. So it could very well be, for all I have said and all I want to say here, that a full and adequate vindication of our moral commitments will vindicate only commitments to consequentialist principles or claims. That is to say that it might turn out that once the relevant values have been identified in terms of which our moral commitments are to be vindicated, and adequate consideration has been given to the extent to which various candidate moral commitments promote the good that is associated with those values, only commitment to, for example, rule-utilitarian principles and the claims such principles licence will be vindicated. What is at stake in this discussion is whether, if that turns out to be the case, it will be the case because the only relevant value in respect of which those consequentialist commitments are vindicated is the value of their instrumental role in action regulation. We have seen, I think, that it would not be for that reason that

they turn out to be adequately vindicated *qua* moral commitments. So the point is that even consequentialist commitments might have – and, I think, are typically supposed to have – value which is not itself brutally consequentialist. If what is being suggested sounds like it requires a deontological conception of morality, it in fact requires only a deontological conception of the *value* of morality, not of its content.

Nobody ought to think, then, that their favoured first-order moral theory requires them to resist the argument which I have been pursuing throughout this chapter that the fully adequate vindication of our moral commitments requires something more than an account of their value in terms of their contribution to the instrumental good of regulating our behaviour. It is natural for us all, then, to ask what is required of a metaethical theory in order for it to furnish us with the resources to undertake such a fully adequate vindication.

What I want to argue is that in order for it to furnish us with the resources to undertake a fully adequate vindication of our moral commitments, a metaethical theory, even a fictionalist theory on the model I have been describing, must be a *realist* theory. I shall explain why even such a fictionalist theory (employing the Character-Centred model of ordinary moral commitment) must be augmented by a realist metaphysics of morality if it is to allow us to adequately vindicate our (on that theory, nondoxastic) moral commitments.

On the Character-Centred model of ordinary moral commitments, those commitments are grounded in our attitudes concerning the sorts of people we want to be (or value being, to put the point in a less neo-Humean way) and our attitudes concerning the extent to which accepting and acting on particular principles and claims is required or permitted by the aim of being that sort of person. It is open to the proponent of a theory employing such a model, as we have just seen, to argue that such commitments are

objectively valuable to the extent that particular commitments of that sort, in fact, play an instrumental role in regulating our behaviour in desirable ways. But as we saw, such an argument is bound to fall short of an adequate vindication of our moral commitments *qua moral*, for it would amount to no more than an argument for the value of such commitments on a par with the value of any of the (non-coercive, choice involving) non-morality-involving mechanisms by which such regulation is or might be achieved. It would, as I said, attempt to vindicate our moral commitments by invoking the wrong sort of value.

But one of the attractions of the Character-Centred picture of ordinary moral commitment is that that picture nicely accommodates a feeling we are liable to have about the character and value of our ordinary morality, namely that in being committed to some moral principle or claim, or in judging something to be wrong, cruel, required-by-decency or whatever, we are seemingly doing something which is more closely connected to the *sort of person we are* in an interesting sense than is what we are doing when accepting or judging other sorts of things. Our moral commitments seemingly tell someone (another person, or ourselves) more about us which is relevant to our *character*, or at least (and this probably better respects the limits of what the seeming data really is) we tend to *value* the having of moral commitments in a way that is tied more closely to the ways in which we judge a person: we are disappointed in our friends when it emerges that they are committed to some moral claim which we find objectionable because it seems to tell us something unwelcome about the sort of person they are, a sort of person we didn't think them to be, or hoped they would not be, or would like them not to be; and similarly we can sometimes find it hard to give appropriate credit to those we dislike for their apparently good moral commitments, for it is an all-too-common fact about us that we sometimes think of those we dislike as objectionable *people*, and it can be

hard to reconcile our judgement of them as objectionable people with their praiseworthy moral commitments precisely because those commitments seem to be some counter-evidence to our judgment of them *as people*. The Character-Centred model of moral commitment affords a ready explanation of this, for on that model our ordinary moral commitments are tied to our attitudes about the sorts of people we value being, and it is clear that a person's attitudes about the sort of person they value being are intimately tied to the sort of person they *are*.

This independently attractive feature of the Character-Centred model can be developed into a way of adequately vindicating our moral commitments. The fact that our moral commitments are reflections of the traits we regard as virtues means that we can make good sense of the thought that there are some moral commitments we ought to have as the thought (or following from the thought) that there are some ways that it is good, *qua* person, to be. For example, we are entitled to regard 'Smacking children is wrong' as something a person ought to accept (as a commitment they ought to have) if we think that that really is a commitment which compassion demands, and we think compassion is a good. Vindicating our moral commitments this way, a way which makes essential use of the link between moral commitment and attitudes which are partly constitutive of character, goes a long way beyond the sort of instrumental vindication which was found to be so inadequate. Unlike the instrumental vindication, the Character-Centred vindication can advert to the non-instrumental good of being a particular sort of person, and to the idea that part of being that sort of person is to be committed to certain moral principles or claims. That non-instrumental vindication can therefore vindicate the thought that there is more to the value of our moral commitments than merely their role in the regulation of behaviour, and that we have reason to think that there are moral commitments we ought to have beyond the interest we have in such

regulation, for we have the additional reason to think that there are some moral commitments we ought to have stemming from the interest we have in people *being* certain sorts of people and the link between being some sorts of people and having some moral commitments.

This train of thought can lead in two directions, each of which requires that we countenance, in our *metaphysics* of morality, that there are not just truths about the character of morality (that it is valuable for such and such reason, for example), but that there are things deserving of the name *moral truths*.

On the first way of developing the vindication of our ordinary moral commitments along the Character-Centred lines just described, the thought is that we are entitled to regard some ways a person might be as good – and therefore that we are entitled to regard those moral commitments tied to that way of being as moral commitments we ought to have – because there *are* virtues in the sense of ways of being (*qua* person) which, as proponents of virtue ethics insist, are in fact admirable, or fine, or good. We might, for example, be drawn to a picture (such as that drawn by Foot (2001)) on which certain character traits, capacities and dispositions to judge are identified as paradigmatically good for the sorts of creatures we are: just as fierceness is good in lions, compassion is good in humans.¹³ Armed with the notion of an objectively good set of ways to be tied to the nature of the sorts of creatures we are, we can help ourselves to the idea that the moral commitments we ought to have are those which are tied in a particularly close way to being *those* ways. On this version of the Character-Centred vindication of our ordinary moral commitments, those ordinary commitments are a matter of our seeing acceptance of some moral claims or action-guiding principles as required by some virtue(s) which we value; and those moral commitments

¹³ It is important that whilst instrumentalists will insist that something is good *for* a lion or a person, the virtue theorist I have in mind insists that something is good *in* a lion or a person.

which we really ought to have (those which make it true that there are some moral commitments we ought to have) are the commitments acceptance of which is *in fact* required by those aspects of our personhood which are *in fact* virtues.

Notice that according to this first way of developing the Character-Centred vindication, the fact that there are some moral commitments we ought to have depends upon a person sometimes being *right* to think that some particular way for them to be is valuable and that acceptance of some claims or principles is required by that valuable way of being. In previous chapters, where our task was to explain what moral commitments *are*, rather than whether they are commitments we really ought to have, it was crucial to see that a person's being committed to some moral claim or principle does not depend upon them being correct in treating acceptance of that claim or principle as required by a particular putative virtue, nor upon their being correct in treating some putative virtue as a virtue. Now, though, we are interested in whether there are any commitments which we ought to have, so it is appropriate to ask whether we would be *right* to treat some putative virtues as virtues, and some claims and principles as required by those virtues. The first version of the vindication of our moral commitments afforded by the Character-Centred account of the *nature* of moral commitments depends upon augmenting that account with a realist account of the existence of the virtues.

The second way of developing the Character-Centred account of the nature of our ordinary moral commitments depends upon a different sort of realism: not realism about the virtues, but realism about the moral facts described or represented by some of the moral claims or principles to which we are (or might be) committed. I noted, in Section 2, that the error theorist is not entitled to answer questions about which moral claims or principles to accept by saying 'the true ones, because they are true'. Indeed, the error

theorist is not entitled to vindicate the thought that there are *any* moral commitments we ought to have by saying that we ought to accept the true ones, because they are true. That is what raised the problem of how to vindicate our moral commitments which the discussion of instrumental vindications was addressing. But of course a theory which allows that there are some true moral claims or principles is under no particular pressure to find an alternative to the simple and intuitive idea that we ought to accept the true moral claims or principles, because they are true.

It might seem that such a realist response to – or perhaps dissolution of – the problem of vindicating our moral commitments faces two immediate problems in the context of our discussion here. Firstly, the idea that we ought to accept the true moral claims or principles, because they are true, might seem to be just as much besides the point of our interest in moral principles (cf. Section 3) as the instrumental vindication turned out to be (in Section 4). Secondly (and perhaps relatedly), it might seem that this could be no proper way to develop a *Character-Centred* vindication, since on the Character-Centred conception of our moral commitments and their role, truth plays no particular part.

Addressing the first worry helps to see why the second is unfounded. It is true – certainly if the hermeneutic strand of the Character-Centred theory of moral commitment is right – that our interest in a person's particular moral commitments often has little or nothing to do with whether those claims or principles to which they are committed are true. That is, as we have seen, at the heart of the motivation for developing a nondoxastic model of ordinary moral commitment. But it does not follow that our interest in whether there are some moral commitments we ought to have is similarly unconcerned with truth. Let me reiterate some important points.

As we saw in Chapter Two, we must be careful to distinguish the theory *of* some practice involving a discourse from that practice itself. Our

ordinary interest in each other's ordinary moral commitments is what motivates the Character-Centred model of our ordinary moral commitments, which is a model of our practice of ordinary moral judgement, deliberation and commitment. But our interest in vindicating the thought that there are some moral commitments we ought to have is a distinctively *unordinary* concern with the status of our moral commitments: it is part of the *theory of* ordinary morality to wonder whether that morality is ever adequately well-founded. The practice of moral judgement and deliberation itself, and the intimately related practice of taking an interest in the morality of ourselves and others, is not the practice of asking questions about the standards of correctness to which moral commitments are subject and whether they meet those standards – they are practices (typically) of *employing*, not *questioning*, such standards.

This distinction between an ordinary practice and the theory of that practice can be obscured if we assume that theorising about some practice is the preserve of philosophers or any other particular sort of person. If we assume that, it will seem implausible to say that the ordinary practice of morality does not involve questioning the standards to which moral commitment (judgment, deliberation) is subject, for clearly ordinary people question such things very often, and if it is assumed that ordinary people are not doing the *theory of* morality when they do it then we must, in order to accommodate that data, say that such questioning is part of the *ordinary practice* of morality, in which we allow that they do participate. But ordinary people theorise (albeit often rather badly) about morality all the time, and so we are under no particular pressure to say that any ordinary foray into questions about whether we are ever properly justified in our moral commitments must be treated as part of the practice of ordinary moral deliberation itself. Likewise, mathematicians might sometimes ask themselves or each other (idly or with great seriousness) about whether their

theories are really, as they have overheard some philosophers of mathematics say, committed to there being abstract objects; but that does not mean that asking questions about ontological commitment to abstract objects is a part of the practice of *mathematics*, rather than a move in the *theory of mathematics* in which ordinary mathematicians are quite free to dabble with whatever degree of seriousness they like.

The theory of morality, then, is interested in (amongst other things) whether or not we can *vindicate* the popular thought that there are some moral commitments we ought to have. The structure I have just reiterated is plain in Mackie's treatment of this issue. Mackie draws a distinction between first- and second-order morality (1977: pp9-10), and proceeds in the second section of his book to show that a particular set of ideas in first-order ethics can be vindicated. But though his discussion in those chapters is mostly of issues in first-order ethics, his purpose is to develop a (framework for) first-order morality which satisfies a *second-order* conception of ethics, namely the instrumental conception discussed already in this chapter. His thought, like mine, is that the vindicatory project involves a second-order concern with the standards by which we ought to measure the propriety of our first-order commitments, and his lengthy discussion of first-order commitments is really an exercise in exploring whether the particular standards identified in his second-order discussion can be met by particular (sorts of) first-order commitments; a second-order exercise in itself.

My criticism of the instrumental vindications offered by Mackie and various fictionalists was that such second-order theories (theories of morality, rather than moral theories) fail to meet their own demands. What the call for a fully adequate vindication demands is not just that *some* good be identified, but that the values attached to our having some commitments in the vindicatory story are those values we attach to our having those commitments already. Mere instrumental role is not the source of the value

which we attach to morality as a distinctive set of commitments, or at least it is not the only source of that value, so any vindication based on just that value is bound to seem less satisfying than one which adverts to a broader range of the things we value about morality. All of these things can be said from within a second-order discussion of the standards of propriety for our moral commitments, a debate in which those standards of propriety are linked (for they must be linked to something, on pain of being merely arbitrary) to what we value.

What drove the Character-Centred model of our ordinary moral commitments was the observation that our *epistemic* relation to the truths of morality, if there are any, is not what we typically assess or value when we think about our own or others' moral commitments, judgements or deliberations. That observation does not automatically undermine the thought that we ought to accept those moral claims or principles which are true, or that we ought to rely on them in moral deliberation. It undermines the thought that our relying on them or accepting them is or ought to be subject to the *epistemic* norms characteristic of *doxastic* acceptance and the employment of doxastically accepted claims or principles. Some argument would be required to show that (a) the fact that the truth of some claims or principles means that we ought to accept them entails that (b) we ought to *believe* those claims and rely on them in deliberation in all and only ways characteristic of doxastically accepted claims or principles. What I want to suggest is that (b) is not true, and that the fact that (b) is not true is at the heart of the Character-Centred model of what the *nature of* our first-order ordinary moral commitments often is and generally ought to be; but that (a) is true, and is at the heart of a second-order vindication of the idea that there are first-order moral commitments *of that sort* (the nondoxastic sort) which we ought to have.

So by way of responding to the worry that suddenly invoking the truth of some moral claims seems to undermine the arguments I proposed in favour of the Character-Centred nondoxastic account of our ordinary moral commitment, I want to say that both the analysis of the nature of what our moral commitments often are and generally ought to be *and* the project of vindicating the thought that there are some moral commitments we ought to have are second-order enterprises. But whilst the analysis of the nature of what our moral commitments often are and generally ought to be is concerned with the ordinary practice of first-order moral judgement and deliberation, the project of vindicating the thought that there are some moral commitments we ought to have is concerned with understanding what our *second-order* ideas about the value of some first-order moral commitments is and assessing whether or not our first order commitments have those features (such as truth, or usefulness) which would make us right to value them in that (second-order) way.

If we have good reason to think that a person really is required by the moral facts to prescind from smacking their child, for example, we have good reason to value being the sort of person for whom smacking their child is not permitted in the way the Character-Centred model of moral commitment describes. That is, we have good reason to value being someone who values compassion and sees smacking their child as incompatible with the demands of being a compassionate person, and as thereby unacceptable. That is to say, I take it, that if we have good reason to think that a person really is required by the moral facts to prescind from smacking their child, we can vindicate the thought that a person ought to be nondoxastically committed in the Character-Centred way to the impermissibility of smacking their child.

Similarly, of course, we can say that if we have good reason to think that a person is required by the facts about social coordination, for example,

to prescind from smacking their child, we have good reason to value being the sort of person for whom smacking their child is not permitted in the way the Character-Centred model of moral commitment describes, and ought to be nondoxastically committed in the Character-Centred way to the impermissibility of smacking their child. As I have said already, the Character-Centred model can play a role in instrumental vindications of our moral commitments just as well as any other account of the nature of moral commitment. The issue, here, is not between accounts of the nature of moral commitment, but between realist and non-realist vindications of the thought that there are some moral commitments we ought to have. Non-realist, instrumental, vindications cannot preserve the distinction between the value of our *moral* commitments and the value of our non-moral commitments in the right sort of way, for if we have good reason to think that a person is required by the facts about social coordination, for example, to prescind from smacking their child, we have just as good reason to value *whatever* promises to control the likelihood of them smacking their child.

It might seem that invoking moral facts gives us no more reason to think that there are some moral commitments we ought to have than invoking facts about social coordination, for example, would give us. But if we think that complying with the moral facts is not, just in itself, as valuable as complying with them *for the right sorts of reasons*, then we are entitled to value being someone who values compassion and sees smacking their child as incompatible with the demands of being a compassionate person, and as thereby unacceptable, as *specially* important, in a way that nobody thinks complying with merely instrumental demands *because they are instrumental demands* is specially important.

This is where the second way of developing the Character-Centred vindication of the thought that there are some moral commitments we ought to have makes contact with the first way, because we are now back in the

realm of *virtue*. The difference between a vindication of that thought based on facts about instrumental role in, for example, social coordination and one based on moral facts which correspond to the claims or principles which are accepted in having some moral commitments is that the facts invoked in the moral realist's vindication uniquely vindicate our having *characteristically moral* attitudes to the facts invoked, for no other sorts of attitudes would be *appropriate* ways of responding to or being sensitive to those facts. Any way of being sensitive to the facts about social coordination would do for the instrumental vindication. But if we are to respond and be sensitive to the moral facts in the right sort of way, we had better have some distinctively moral attitudes, attitudes which involve concern for compassion, integrity and the other ways of being – virtues – which have special importance for us in virtue of their contributing to a person's moral character. Responding to the moral facts by means of attitudes other than these is bound to seem like responding in the wrong sort of way, and our consequent actions to seem like the right actions done for the wrong sorts of reasons, even if they are the result of *some* sort of sensitivity to the moral facts.

This might be what we are trying to articulate when we think of morality as tied in a particularly close way to character. And it might be what makes the realist Character-Centred vindication of our *nondoxastic* moral commitments more satisfying in respect of its adverting to the things we really value about morality than either merely instrumentalist vindications of nondoxastic morality (which have no place for the idea that there are moral truths which determine the particular reasons for which a person ought to act) or traditional realist vindications of doxastic morality which put no special emphasis on any feature of our moral commitments other than their brute veridicality. The difference between the first and second ways of working out the Character-Centred idea of moral commitment into a vindication of the idea that there are some moral

commitments we ought to have is that on the first (virtue-theory) way of working out that idea it is a fact about us that there are some ways it is good to be, some virtues we ought to manifest, whilst on the second version of the idea, it is a fact about the moral truths that there are that they make it the case that we ought to be some ways, in order to respond to those moral truths appropriately.

6. Conclusion

So much for two of the ways in which the Character-Centred model of our ordinary moral commitment might be combined with a realist theory of either the virtues or of the truth of some of the particular claims or principles which, in having some moral commitments, we accept. The two ways I have discussed do not exhaust the options. On a view of the relation between moral facts and the virtues such as that of Wiggins, [1987], for example, the gap between the virtue theory version of the realistic fictionalist's vindication of our moral commitments and the second, moral fact centred, version is not even as wide as the slightly attenuated gap I have insisted upon. No matter. My concern here has been to show that there are options. They are options which we may want to avail ourselves of *if* we are concerned to adequately vindicate the thought that there are some moral commitments we ought to have. They are options which we may feel are ruled out by our other metaethical commitments. But the model is at least internally coherent, and it does justice to some important data about our prevailing attitudes to the value, function and nature of our moral commitments, and deserves, in virtue of that, to be taken seriously as one option which has not adequately been entertained or explored because of a too widespread conviction that fictionalism must, if it is to vindicate our ordinary moral commitments at all, find ways of doing so without recourse

to the natural and appealing idea that there are some moral commitments we ought to have *because the moral facts are as they are*.

– Conclusion –

Conclusion

Fictionalism can be very useful for realists, and realism can be very useful for fictionalists; and realistic fictionalism is a perfectly coherent position. I have tried, in Part II, to show that realistic fictionalism about our ordinary moral commitments is a well motivated and defensible position. But the lesson of Part I is that whether we are working on metaethics, the philosophy of mathematics, the philosophy of science, metaphysics or any branch of philosophy in which realism is an option and fictionalism seems a viable way of resisting the realists' traditional emphasis on belief and knowledge, we ought to be open to the possibilities afforded by realistic fictionalism.

As it happens, I am rather drawn to the idea that realistic fictionalism might be a good theory in the philosophy of fiction (though it was one of the lessons of Part I that *fictionalism* is not tied to any theory of fiction), as well as in the philosophy of our ordinary moral commitments. Working that out is work for elsewhere, of course. And I certainly do *not* want to leave the impression that I think realistic fictionalism might be a good theory in all or even most of the fields in which fictionalism and realism have traditionally been opposed. (I do not, for example, think that realistic fictionalism about mathematics is a particularly appealing position, because I don't think any sort of fictionalism about mathematics is particularly appealing, though I offer no defence of that view here.) My purpose, in so far as general lessons about fictionalism are concerned, has been to dispel a common (if not universal) assumption that, at best, there could be no good motivation for realistic fictionalism and, at worst, that such a beast would be a figment of a confused and incoherent imagination. So far as the metaethics developed in Part II is concerned, I have made a start on what I take to be a promising and

potentially very fruitful project of understanding our moral commitments, judgements and deliberation using a nondoxastic model which leaves adequate room for a fully realist metaphysics of morals. For those of us, at least, who are inclined to be moral realists, but who harbour serious doubts about the realists' traditional accounts of our moral attitudes, that seems worthwhile.

Of course, many issues remain outstanding. Here are just a few.

The theory of acceptance I proposed in Chapter Three is, I think, very likely to be broadly right; it is respectful of our intuitive judgements about what it takes for someone to 'accept' something (in the relevant range of senses), at least. But there is far more to be said about the way in which such an 'internalist' theory, which adverts only to the norms *treated as* applying, relates to the sorts of normativist theories (which I am bound to call 'externalist') which distinguish belief from other mental states (completely or in part) in virtue of the norms which *do* apply to the various attitudes. Such normativism is accepted by, for example, Bratman, 1987, Harman, 1986 and Kalderon, 2005a to name just a few of those already cited in this thesis, and is at least widely accepted if it is not the prevailing view. I do not want to say that normativism is completely wrong, that there is no interesting and quite deep connection between differences between attitudes and the norms which, in fact, apply to them. Indeed, at one point I thought that normativism is the correct way of making the relevant distinctions, and I see my move towards a more internalist picture as an attempt to retain what I can of the normativists' insight whilst keeping hold of the crucial idea that the characterisation of a person's mental or psychological states must be prior to the question of which norms apply to things (mental or psychological states) of that sort. It would not surprise me at all if there turns out, on further investigation, to be an essential connection between the norms which, in fact, apply to some mental states and the norms invoked *as*

applying in being in those mental states, in which case I would be happy to regard normativism as a perfectly appropriate way of characterising the differences between, for example, belief and nondoxastic acceptance, but as less fundamental an explanation of that difference than the theory of acceptance I have proposed. But all of that must await further work on the characterisation of mental (or psychological) states such as acceptance in terms of norms.

Something which rather surprised me as I wrote Chapter Six was how close to virtue theory I was coming. I have never been at all attracted to the (sterio?)typical virtue theorists' insistence on the primacy of character in moral evaluation. Talk of character in the Character-Centred theory of ordinary moral commitment is, as I discussed in Chapter Four, not by any means tantamount to acceptance of a virtue theoretic account of the demands of morality, or of what the *content* of a moral theory or set of commitments ought to be: talk of character in the Character-Centred theory of ordinary moral commitment is meant to involve character just in the theory of moral *commitment*, not the theory of which commitments are *good*. But as I pointed out in Chapter Six, a full-blown virtue theory which delivers conclusions about which commitments are good, and explains that in ways which put character virtues at the heart of the matter, *might* quite intelligible be appended to the Character-Centred theory of ordinary moral commitment in order to vindicate the thought that some of our moral commitments are the right ones to have after all. What rather surprised me was the extent to which I found myself linking the *other* way of introducing realism into the Character-Centred theory (or appending realism to it) with character in what might seem to some a decidedly virtue theoretic way. I ended up, you will recall, saying that a person who thinks that some particular moral facts obtain might well want to say that a person ought to be a particular sort of person *in order to respond to those facts in an (or the) appropriate way*, an idea

which I linked to the appealing thought that there is something distinctively important about acting in the right way *for the right reasons*. Talk of acting for the right reasons and the relation of that to being a particular sort of person deserves a great deal more exploration, but it is striking that on the understanding of these things which I offered to the fictionalist as a vindication of the thought that there are some moral commitments we ought to have, we appear to end up in much the same place as the virtue theorist, arguing that there is some *sort of person* we ought to be. As I said towards the end of Chapter Six, I think that is a conclusion to which all sorts of realist are entitled which doesn't (or oughtn't to) count as embracing virtue ethics as such, for no claim to the *primacy* (in any interesting sense) of character is involved. It would repay further examining whether a realistic fictionalism can do without even that much appeal to anything like *virtue*, though.

Incidentally, and here is an issue which might appear to be outstanding but which I do not *think* is, the sense of 'character' employed throughout *my* discussion of character in the context of the Character-Centred model and theory is not contentious enough to be subject to much further debate about the very existence of such a thing as a person's 'character'. There are those who doubt that (what I'll call) *character*, a stable, underlying state of a person which is apt to explain our actions and choices (for example), exists – we speak of a person's *character* but in fact there are just a person's reactions to situations they find themselves in, which sometimes display some regularity but which, *if* they do, are not explained by any underlying state with special very general responsibility for our choices and attitudes. Whatever the merits of that view, I do not mean to be speaking of *character*, in the contested sense, but of character in a more colloquial sense. Everything I say about character can, I think, be read as referring to something *impermanent*, not particularly stable and certainly not metaphysically expensive.

The topic of the nondoxastic and its roles in our thinking has only begun to be explored by the discussion in this thesis. Interesting avenues such as the relation between my nondoxastic picture and Robert Adams's idea that there is a role for what he calls 'moral faith' (Adams, 1995), and the relation of both ideas to the well-known Kantian doctrine of 'rational faith' (in, for example, Kant, [1794]) remain to be explored (indeed, Kant, [1786] is a very nice discussion of rational faith compared to rational belief). Neither Adams nor Kant have in mind what counts as the nondoxastic on my understanding of that, but proper consideration of their views (and the views of others who have addressed similar issues in sympathetic ways) is bound to shed light on the problems I have been concerned with, and might very reasonably be expected to suggest further applications of the ideas I have begun to develop.

No doubt many other issues remain to be addressed too. Certainly, realistic fictionalism has not, for all I've done in this thesis, been *proved* by any means. But I hope that I have done enough to have earned the right to say, to those who are inclined to think of the realistic fictionalist as necessarily confused, as the man that has something bountifully laughable about him, that there is more in that man than perhaps they think for.

References

In lieu of an index, the reader will find references to chapters (in bold) and section numbers in which works are cited following each entry in square brackets.

Robert Merrihew Adams, 1995. 'Moral Faith', *The Journal of Philosophy* vol.92 no.2 [Conc.]

Julia Annas & Jonathan Barnes, 1985. *The Modes of Scepticism: Ancient Texts and Modern Interpretations*, Cambridge: CUP [5,4]

Jamin Asay, forthcoming. 'Truthmaking, Metaethics, and Creeping Minimalism', *Philosophical Studies* [Intro.,2]

Robert Audi, 2007. 'Belief, Intention, and Reasons for Action' in Timmons *et al*, 2007 [5,4]

A. J. Ayer, [1936/46] 1971. *Language, Truth and Logic*, Harmondsworth: Penguin [Intro.,2; 2,2; 5,4]

Michael Beaney (ed.), 1997. *The Frege Reader*, Oxford: Blackwell

Paul Benacerraf, [1973] 1983. 'Mathematical Truth', *Journal of Philosophy* 70, reprinted in *Philosophy of Mathematics: Selected Readings (2nd Edition)*, eds. Paul Benacerraf & Hilary Putnam, Cambridge: CUP [2,3]

Jonathan Bennett, 1990. 'Why Are Beliefs Involuntary?', *Analysis* 50:2 [3,2]

Simon Blackburn, 1984. *Spreading the Word*, Oxford: OUP [1,2; 2,2; 3,6; 5,4]

_____, [1986]. 'Morals and Modals', reprinted in Blackburn, 1993 [5,4]

_____, 1993. *Essays in Quasi-Realism*, Oxford: OUP [Intro.,2; 2,2; 3,6]

_____, 1998. *Ruling Passions: A Theory of Practical Reasoning*, Oxford: OUP [Intro.,2]

_____, 2005. 'Quasi-Realism No Fictionalism' in Kalderon 2005a [Intro.,2; Intro.,3; 1,2]

Richard Boyd, 1990. 'Realism, Conventionality, and "Realism About"', in *Meaning and Method: Essays in Honour of Hilary Putnam* ed. George Boolos, Cambridge: CUP [Intro.,2]

Michael Bratman, 1987. *Intention, Plans, and Practical Reason*, Cambridge, Mass: Harvard University Press [3,3; Conc.]

- Stuart Brock**, 2002. 'Fictionalism about Fictional Characters', *Noûs* 36:1 [Intro.,1; 1,5; 2,5]
- Jessica Brown**, 2008. "Subject-Sensitive Invariantism and the Knowledge Norm for Practical Reasoning." *Nous* 42 [5,4]
- L. Jonathan Cohen**, 1989. 'Belief and Acceptance', *Mind* vol.98 no.391 [3,2]
- _____, 1992. *An Essay on Belief and Acceptance*, Oxford: OUP [3,2]
- Alice Crary**, 2007. *Beyond Moral Judgement*, Cambridge, Mass: Harvard University Press [4,4]
- Arthur Conan Doyle**, [1902] 2010. *The Hound of the Baskervilles*, London: Collins Classics [1,1]
- David Copp**, 2001. 'Realist-Expressivism: A Neglected Option for Moral Realism', *Social Philosophy and Policy* 18 [Intro.,2; 3,1]
- Roger Crisp**, 2007. 'Intuitionism and Disagreement', in Timmons et al, 2007 [5,4]
- Chris John Daly**, 2008. 'Fictionalism and the Attitudes', *Philosophical Studies* 139 [1,3; 2,1; 2,2; 2,3 3,1]
- Stephen Darwall**, 2006. *The Second-Person Standpoint: Morality, Respect, and Accountability*, Cambridge, Mass: Harvard University Press [4,3]
- Michael Devitt**, 1991. *Realism and Truth*, 2nd Edition, Princeton: Princeton University Press [Intro.,2]
- John Dewey**, 1945. 'Ethical Subject-Matter and Language', *The Journal of Philosophy* vol.42 no.26 [Intro.,2]
- Fyodor Dostoevsky**, [1866] 2000. *Crime and Punishment*, trans. Constance Garnett, Ware: Wordsworth Classics [4,1; 4,4]
- James Dreier**, 2004. 'Meta-Ethics and the Problem of Creeping Minimalism', *Philosophical Perspectives* 18 [Intro.,2]
- Michael Dummett**, [1964] 1978. 'Realism', in *Truth and Other Enigmas*, London: Duckworth [Intro.,2]
- _____, [1982]. 'Realism', in Dummett, 1993 [Intro.,2]
- _____, [1992]. 'Realism and Anti-Realism', in Dummett, 1993 [Intro.,2]

- _____, 1993. *The Seas of Language*, Oxford: OUP
- Matti Eklund**, 2005. 'Fiction, Indifference, and Ontology', *Philosophy and Phenomenological Research* vol.71 no.3 [4,1]
- _____, 2009. 'Fictionalism', *The Stanford Encyclopedia of Philosophy* (Spring 2009 Edition), Edward N. Zalta (ed.) <http://plato.stanford.edu/archives/spr2009/entries/fictionalism/> [1,3]
- Pascal Engel** (ed.), 2000. *Believing and Accepting*, Dordrecht: Kluwer [3,2]
- Richard Feldman & Ted A. Warfield** (eds.), 2010. *Disagreement*, Oxford: OUP [5,4]
- Hartry Field**, 1980. *Science without Numbers: A Defense of Nominalism*, Oxford: Blackwell. [Intro.,1; Intro.,2; 1,3; 2,3; 2,5; 3,5; 6,3; 6,4]
- _____, 1989. *Realism, Mathematics & Modality*, Oxford: Blackwell [1,3; 2,3; 2,5; 3,5; 6,3]
- Philippa Foot**, 2001. *Natural Goodness*, Oxford: OUP [4,3; 6,5]
- Bas van Fraassen**, 1980. *The Scientific Image*, Oxford: OUP [Intro.,1; 1,1; 1,2; 2,2; 2,3; 2,5; 3,3; 3,5; 5,4]
- Gottlob Frege**, [1897]. 'Logic', in Beaney, 1997 [1,4]
- _____, [1918]. 'Thought', in Beaney, 1997 [3,4]
- Stacie Friend**, 2007. 'Fictional Characters', *Philosophy Compass* 2 [1,4]
- Peter Geach**, 1965. 'Assertion', *The Philosophical Review* vol.74 no.4 [Intro.,2]
- Allan Gibbard**, 1990. *Wise Choices, Apt Feelings*, Cambridge, Mass: Harvard University Press [2,2; 3,3; 5,4]
- _____, 2003. *Thinking How to Live*, Cambridge, Mass: Harvard University Press [Intro.,2; Intro.,3]
- Carl Ginet**, 2001. 'Deciding to Believe', in *Knowledge, Truth, and Duty: Essays on Epistemic Justification, Responsibility, and Virtue*, ed. Matthias Steup, Oxford: OUP [3,2]
- Christopher W. Gowans**, 2000. *Moral Disagreements: Classic and Contemporary Readings*, London: Routledge [5,4]
- Patrick Greenough & Michael P. Lynch**, 2006. *Truth and Realism*, Oxford: OUP

- R. M. Hare**, 1952. *The Language of Morals*, Oxford: Clarendon [2,2; 5,4]
- _____, 1997. *Sorting Out Ethics*, Oxford: OUP [5,4]
- Gilbert Harman**, 1986. *Change in View*, Cambridge, Mass: MIT Press [1,2; 5,4; **Conc.**]
- John Hawthorne**, 2004. *Knowledge and Lotteries*, Oxford: Clarendon [5,4]
- John Hawthorne & Jason Stanley**, 2008. 'Knowledge and Action', *Journal of Philosophy* vol.CV no.10 [5,4]
- Paul Horwich**, [1991] 2004. 'On the Nature and Norms of Theoretical Commitment', reprinted in his *From a Deflationary Point of View*, Oxford: OUP [2,1; 3,4]
- _____, 1998. *Truth*, 2nd edition, Oxford: OUP [**Intro.**,3; 3,4]
- _____, 2006. 'A World without Isms', in Greenough & Lynch, 2006 [2,1; 3,4]
- David Hume**, [1752] 1987. 'Of Moral Prejudices', in *Essays, Moral, Political, and Literary, Part II*, reprinted in *Essays Moral, Political, and Literary* (Revised Edition), ed. Eugene F. Miller, Indianapolis: Liberty Fund [5,1]
- Robert N. Johnson**, 2003. 'Virtue and Right', *Ethics* vol.113 no.4 [4,2]
- Richard Joyce**, 2001. *The Myth of Morality*, Cambridge: CUP [**Intro.**,1; 2,5; 3,6; 6,2; 6,4]
- _____, 2005. 'Moral Fictionalism' in Kalderon, 2005b [1,3; 2,5; 3,3; 3,6; 6,2; 6,4]
- _____, 2006. *The Evolution of Morality*, Cambridge, Mass: MIT Press [6,2; 6,4]
- Mark Eli Kalderon**, 2005a. *Moral Fictionalism*, Oxford: OUP [**Intro.**,1; **Intro.**,2; 1,2; 2,2; 3,1; 3,3; 3,6; 5,4; **Conc.**]
- _____, (ed.), 2005b. *Fictionalism in Metaphysics*, Oxford: OUP
- _____, 2005c. 'Introduction', in Kalderon, 2005b [**Intro.**,2; 2,2]
- Immanuel Kant**, [1785] 1996. *Groundwork of the Metaphysics of Morals*, in *Practical Philosophy*, trans. and ed. Mary J. Gregor, Cambridge: CUP [6,5]
- _____, [1786]. 'What does it Mean to Orient Oneself in Thinking?', in Wood & Giovanni, 1998 [**Conc.**]
- _____, [1794]. *Religion within the Boundaries of Mere Reason*, in Wood & Giovanni, 1998 [**Conc.**]

- Jerrold J. Katz**, 1998. *Realistic Rationalism*, Cambridge, Mass: MIT Press [Intro.,2]
- Frederick Kroon**, 2001. 'The Semantics of 'Things in Themselves': A Deflationary Account', *The Philosophical Quarterly*, vol.51 no.203 [Intro.,1]
- Mary Leng**, 2010. *Mathematics and Reality*, Oxford: OUP [1,2]
- Alan Leslie**, 1994. 'Pretending and Believing: Issues in the Theory of ToMM', *Cognition* 50 [2,5]
- David Lewis**, 1978. 'Truth in Fiction', *American Philosophical Quarterly* 15:1 [1,4]
- J. L. Mackie**, 1977. *Ethics: Inventing Right and Wrong*, Harmondsworth: Penguin [Intro.,3; 6,2; 6,4; 6,5]
- John McDowell**, 1994. *Mind and World*, Cambridge, Mass: Harvard University Press [3,3]
- Herman Melville**, [1851] 2007. *Moby Dick, or The Whale*, London: Vintage Books [2,5]
- Stephen Neale**, 2001. *Facing Facts*, Oxford: OUP [Intro.,2; 1,3]
- Ram Neta**, 2009. 'Treating Something as a Reason for Action', *Noûs* 43:4 [5,4]
- Shaun Nichols**, 2004. 'Imagining and Believing: The Promise of a Single Code', *The Journal of Aesthetics and Art Criticism* 62:2 [2,5]
- Daniel Nolan**, 2005. 'Fictionalist Attitudes about Fictional Matters', in Kalderon, 2005b [Intro.,1; Intro.,3]
- Daniel Nolan, Greg Restall & Caroline West**, 2005. 'Moral Fictionalism Versus the Rest', *Australasian Journal of Philosophy*, 83:3 [1,3]
- Robert Nozick**, 1974. *Anarchy, State, and Utopia*, New York: Basic Books [6,4]
- David Papineau**, 1988. 'Mathematical Fictionalism', *International Studies in the Philosophy of Science* vol.2 no.2 [2,5]
- _____, 1993. *Philosophical Naturalism*, Oxford: Blackwell [3,1; 5,4]
- Hilary Putnam**, 1994. 'Pragmatism and Moral Objectivity', in his *Words and Life*, ed. James Conant, Cambridge, Mass: Harvard University Press [6,2]
- _____, 2002. *The Collapse of the Fact/Value Dichotomy, and Other Essays*, Cambridge, Mass: Harvard University Press [6,2]

- _____, 2004. *Ethics without Ontology*, Cambridge, Mass: Harvard University Press [3,6]
- Willard V. Quine**, 1948. 'On What There Is', *The Review of Metaphysics* vol.2 no.5 [1,3]
- D. D. Raphael**, 1983. 'Can Literature Be Moral Philosophy?', *New Literary History* 15 [4,4]
- Joseph Raz**, 2006. 'The Problem of Authority', *Minnesota Law Review* vol.90 no4 (April) [5,3]
- Richard Rorty**, [1985] 1991. 'Solidarity or Objectivity?', reprinted in his *Objectivity, Relativism, and Truth: Philosophical Papers vol.1*, Cambridge: CUP [6,2]
- _____, 1989. *Contingency, Irony, and Solidarity*, Cambridge: CUP [6,2]
- Gideon Rosen**, 1990. 'Modal Fictionalism', *Mind* vol.99 no.395 [Intro.,1; 1,3; 3,5]
- _____, 1994. 'Objectivity and Modern Idealism: What is the Question?', in *Philosophy in Mind* eds. Michaelis Michael and John O'Leary-Hawthorne, Dordrecht: Kluwer [Intro.,2; 2,2]
- Mark Sainsbury**, 2009. *Fiction and Fictionalism*, Oxford: Routledge [1,1; 1,3; 1,4; 3,7; 5,4; 5,5]
- Robert K. Shope**, 1978. 'The Conditional Fallacy in Contemporary Philosophy', *Journal of Philosophy* vol.75 no.8 [4,2]
- Mark Schroeder**, 2008. *Being For: Evaluating the Semantic Program of Expressivism*, Oxford: OUP [Intro.,2]
- Henry Sidgwick**, [1874/1907] 1981. *The Methods of Ethics* 7th Edition, Cambridge: Hackett [5,4]
- Jason Stanley**, 2001. 'Hermeneutic Fictionalism', *Midwest Studies in Philosophy* 25 [1,5; 4,1]
- Charles L. Stevenson**, 1944. *Ethics and Language*, New Haven: Yale University Press [Intro.,2; Intro.,3; 2,2; 5,4]
- _____, [1948]. 'The Nature of Ethical Disagreement', reprinted in Stevenson, 1963a [5,4]
- _____, 1963a. *Facts and Values: Studies in Ethical Analysis*, New Haven: Yale University Press

- _____, 1963b. 'Retrospective Comments' in Stevenson, 1963a
[Intro.,3; 2,2; 3,6]
- Kathleen Stock**, 2011. 'Fiction and Imagination I: Fictive Utterance and Imagining',
Proceeding of the Aristotelian Society vol. LXXXV [1,2; 2,5]
- P. F. Strawson**, [1962] 2008. 'Freedom and Resentment', reprinted in his *Freedom and Resentment and Other Essays*, Oxford: Routledge [4,3]
- Amie L. Thomasson**, 2007. *Ordinary Objects*, Oxford: OUP [6,2]
- Mark Timmons, John Greco & Alfred R. Mele** (eds.), 2007. *Rationality and the Good: Critical Essays on the Ethics and Epistemology of Robert Audi*, Oxford: OUP
- Charles Travis**, 2006. *Thought's Footing: A Theme in Wittgenstein's Philosophical Investigations*, Oxford: OUP [3,3]
- J. David Velleman**, 2001. 'The Genesis of Shame', *Philosophy and Public Affairs* vol.30 no.1 [4,3]
- Gerald Vision**, 1994. 'Fiction and Fictionalist Reductions', *Pacific Philosophical Quarterly*, 74 [1,3]
- Kendall Walton**, 1990. *Mimesis As Make Believe: On the Foundations of the Representational Arts*, Cambridge, Mass: Harvard University Press [1,1]
- Ralph Wedgwood**, 2007. *The Nature of Normativity*, Oxford: OUP [4,3]
- _____, 2010. 'The Moral Evil Demons', in Feldman & Warfield, 2010 [5,4]
- David Wiggins**, [1987] 2002. 'A Sensible Subjectivism?', in his *Needs, Values, Truth*, 3rd Edition, Oxford: OUP [6,6]
- _____, 2006. *Ethics: Twelve Lectures on the Philosophy of Morality*, London: Penguin [1,2]
- Oscar Wilde**, [1891] 2010. *The Picture of Dorian Gray*, London: Collins Classics [5,5]
- Bernard Williams**, 1993. *Shame and Necessity*, Berkley: University of California Press [4,3]
- Timothy Williamson**, 2000. *Knowledge and its Limits*, Oxford: OUP [3,3]
- _____, 2005. 'Contextualism, Subject Sensitive Invariantism, and Knowledge of Knowledge', *Philosophical Quarterly* 55 [5,4]
- _____, 2006. 'Must Do Better', in Greenough & Lynch, 2006 [Intro,2]

Ludwig Wittgenstein, [1921] 1974. *Tractatus Logico-Philosophicus*, Oxford: Routledge [Intro.,2]

_____, 1978. *Remarks on the Foundations of Mathematics*, Cambridge, Mass: MIT Press [1,1]

Jonathan Wolff, 2002. *Why Read Marx Today?*, Oxford: OUP [3,5]

Allen Wood & George Di Giovanni, (eds.) 1998. *Immanuel Kant: Religion within the Boundaries of Mere Reason, And Other Writings*, Cambridge: CUP

Crispin Wright, 1980. *Wittgenstein on the Foundations of Mathematics*, London: Duckworth [1,1]

_____, 1992. *Truth and Objectivity*, Harvard: Harvard University Press [1,2]

_____, 1995. 'Truth in Ethics', *Ratio* (New Series) VIII no.3 (Dec) [Intro.,3]

_____, 1999. 'Truth: A Traditional Debate Reviewed', reprinted in *Truth*, eds. Simon Blackburn and Keith Simmons, Oxford: OUP [3,3]

Stephen Yablo, 2000. 'Apriority and Existence' in *New Essays on the A Priori* eds. Paul Boghossian & Christopher Peacocke, Oxford: OUP [Intro.,1; 2,3]

_____, 2002. 'Go Figure: A Path through Fictionalism', *Midwest Studies in Philosophy*, 25 [Intro.,1; 2,1; 2,4]

_____, 2005. 'The Myth of the Seven' in Kalderon 2005b [Intro.,1; 2,5; 6,3]

Edward N. Zalta, 1988. *Intensional Logic and the Metaphysics of Intentionality*, Cambridge, Mass: MIT Press [3,7]