University of Reading

A Metaethic Made Useful

A Wittgenstein-Inspired Approach to Metaethics

Thesis submitted for the degree of Doctor of Philosophy
Department of Philosophy
David Mutch
October 2016

Declaration

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

David Mutch
24th January 2017

Abstract

In this thesis, I use Wittgenstein's thoughts about ethics, and about philosophy in general, to argue in defence of a form of metaethical subjectivism. I argue that a Wittgensteinian metaethic can have value not only as a 'therapy' for philosophical unease, nor simply because it can help us to find the truth, but also because conceptual confusions can impact upon substantive moral thinking. Wittgenstein's thoughts on ethics, I argue, indicate a broadly anti-realist position, and many bear a striking resemblance to later non-cognitivist qualms about simple subjectivism. His expressivism about psychological avowals, in particular his comments on psychological self-*descriptions*, allow for a nuanced form of expressivism, and may provide a model for ethical language. Whilst the contemporary expressivist's 'Parity Thesis' helps expressivism to avoid collapse into simple subjectivism, it also highlights the need to tackle 'Moore's Paradox'. Wittgenstein's solution to the paradox, however, indicates that no story about the meanings of moral sentences can follow from the non-cognitivist's story about the meanings of simple moral utterances. In preference to non-cognitivism, then, I argue, we should return to a more simple-minded subjectivism, the main objections to which, viz. the disagreement and modal problems, can be overcome by Wittgenstein's accounts of meaning and truth. I argue that 'idealizing subjectivism' is not objectionably *ad hoc*, since the aim of 'idealizing' is correctly to capture the use of moral terms, which can be the only criterion for correctness for a philosophical analysis, and that, as social creatures and people of conscience, we have a deep need for the moral concepts the idealizing subjectivist describes. Finally, I argue, with reference to Wittgenstein's thoughts on ritual practices, that there are more constraints on legitimate moral judgement than are typically allowed by anti-realists, but not the kinds of constraint typically claimed by their opponents.

For Sophie and for my parents.

With thanks to Luke for getting me started, to Simon for getting me finished, to Bart, Brad and Severin for your wisdom, and to my peers in the department for a thousand conversations.

"In the train of every idea that costs a lot come a host of cheap ones: amongst them even a few that are useful."
(CV, p. 66)

Contents

Introduction

Wittgenstein once remarked that many philosophical problems arise due to the philosopher's "craving for generality" and his "contemptuous attitude towards the particular case" (BB, p. 18). It may seem somewhat ironic, then, that, throughout his lifetime, Wittgenstein was continually drawn towards questions which seem to be of fundamental, and, indeed, extremely general import.

Wittgenstein's academic career began not in philosophy, but in engineering. His research into novel forms of aeronautic propulsion, however, soon led him to an interest in pure mathematics. Almost as quickly, his interest in mathematics itself gave way to an absorption in the philosophy of mathematics. It was this subject that was the initial focus of his studies under Bertrand Russell in Cambridge.

The questions that had initially piqued Wittgenstein's interest in philosophy, however, occupied only a relatively minor place in his thinking by the time *Tractatus Logico-Philosophicus* was completed (Schroeder, S. 2006a, pp. 14-20). Indeed, in the preface to that book, Wittgenstein claimed not only to have solved the problems which first attracted him to philosophy, but to have solved, in essence, *all* the problems of philosophy (TLP, pp. 3f).

Whilst Wittgenstein later came to recognise "grave mistakes" (PI, p. viii) in his youthful claims, the pursuit of something fundamental in philosophy which characterised his early philosophical development was never repudiated, and whilst he never repeated his hubristic claim to have solved all philosophical problems, he nonetheless maintained that he had found a 'method' for doing so (PI §133; Moore 1955, p. 26; CV, p. 3).

Whilst Wittgenstein considered the development of his general philosophical

'method'[1] his most important contribution to philosophy (Moore 1955, p. 26; CV, p. 3), he did not confine his philosophical investigations only to such rarefied concerns. Indeed, he thought that his method could only be demonstrated by examples of its application to specific cases (PI §133).

In the preface to *Philosophical Investigations*, Wittgenstein lists amongst the book's topics, problems concerning

> ...the concepts of meaning, of understanding, of proposition, of logic, the foundations of mathematics, states of consciousness, and other things. (PI, p. vii)

In his lifetime, Wittgenstein made significant contributions to areas as disparate as the philosophies of mind, science, mathematics, aesthetics, religion, action and epistemology.[2] There is, however, a notable and somewhat puzzling lacuna in Wittgenstein's efforts to apply his own philosophical method, namely, in the arena of ethics.

Unanimity of opinion is a rare commodity in the marketplace of Wittgenstein scholarship. Notwithstanding, there is one item which is almost universally on offer when stalls on Wittgenstein's ethics are set out. The curio in question is the observation that, amongst the great philosophers, Wittgenstein was exceptionally (if not uniquely (Schroeder, S. 2006a, p. 99)) preoccupied with the ethical aspect of life (Monk 1990, p. 278; Glock 2015, p. 99).

This is not to say that he thought about such matters in exceptional philosophical depth. Indeed, customarily displayed adjacent to the aforementioned curio is the further

1 Whilst Wittgenstein made repeated references to his philosophical 'method', he also recognised that such a characterisation was misleading. There is not *a* philosophical method, but, rather, there are *various* methods, all sharing an overriding *aim*, namely, distinguishing between sense and nonsense in order to clear up philosophical confusion (PI §133; Baker & Hacker 1980a, p. 290).

2 This is by no means an exhaustive catalogue, but serves only as an indication of the breadth of Wittgenstein's philosophical concerns.

observation that, in contra-position to his uncommon ethical concern, Wittgenstein's recorded contributions to moral philosophy are notable largely due to their absence. Wittgenstein's preoccupation with ethics rarely manifested itself in philosophical inquisitiveness; rather, it was of an extraordinarily deeply felt and personal nature. Wittgenstein was obsessed with his own moral character,[3] and, throughout his life, he was racked by feelings of moral inadequacy – feelings which, on numerous occasions, drove him close to suicide (McGuinness 1988, p. 154; Schroeder, S. 2006a, p. 102).

It isn't, however, that Wittgenstein thought that there was nothing worth saying on the subject. Although his treatment of ethics in the *Tractatus* was, to say the least, breviloquent, he nonetheless attached great importance to it (Schroeder, S. 2006a, pp. 29f), once going so far as to claim (somewhat hyperbolically) that "the point of the book is ethical" (LF, p. 94). In his 'Lecture on Ethics', he characterised ethics as "a subject which seems to me to be of general importance", and stressed that it was one on which he was "keen on communicating" his thoughts (LE, p. 4). Again, in a lecture in 1933, he remarked that he had "always wanted to say something about the grammar of ethical expressions" (Moore 1955, p. 16).

As G.E. Moore (1955, p. 16) notes, though, on this last occasion Wittgenstein in fact went on to say very little specifically about ethics, preferring to tackle the subject mainly tangentially via discussions of aesthetics, religion, and ritual practices, which he took to illustrate salient points. This oblique approach was characteristic of Wittgenstein's later contributions to ethics, of which what we have is to be found largely in lectures ostensibly on the subject of aesthetics (LC), with only passing gestures at the relevance of such discussions to ethics. Perhaps the most direct discussion of ethics in Wittgenstein's later period is found, not in his notes or lectures, but in records of conversations (Rhees

3 Perhaps even to the extent that the vanity involved in such an obsession itself became a moral blemish (Schroeder, S. 2006a, p. 101).

1965, pp. 21-4), and in his major work, *Philosophical Investigations*, ethics is mentioned only in two remarks (PI §§77, 304).

As a consequence of Wittgenstein's relative silence on ethics, his influence on the subject is rarely detected. Some have claimed that, in the case of first-order normative and applied ethics at least, this apparent absence is not much to be lamented.

For one thing, his indubitable strength of feeling does not seem to have made him a paragon of virtue, at least not in an Aristotelian sense. Wittgenstein was notoriously intemperate and uncompromising, often seeming to have scant regard for the effects of his intense personality on those around him (Glock 2015, p. 100).

His occasionally difficult behaviour, though, seems to have been symptomatic not of a malign will, bent on causing distress, but of the relentless pursuit of commendable ideals of honesty and personal integrity. As Bertrand Russell once remarked:

> His nature is good through and through...[H]e might do all sorts of things in a passion, but he would not practise any cold-blooded immorality. (Monk 1990, p. 52)

Having said this, some of Wittgenstein's ideals, too, have been brought into question, for what we have of Wittgenstein's personal opinions (recorded largely only in diary entries and records of conversations, and, it should be stressed, never intended for publication (Glock 2015, p. 100)) has struck some as at least morally ambiguous.

Although Jewish himself by ancestry,[4] some of his remarks have been interpreted as suggestive of anti-Semitic tendencies (Monk 191, p. 314; Glock 2015, p. 100). Others have appeared to some to display a misogynistic bent (Glock 2015, p. 100). He appears to have been inclined towards social conservatism and religious and political authoritarianism

4 His family converted to Christianity before his birth (Monk 1991, pp. 4f).

(Glock 2015, p. 100), expressing approval of the kind of strict class-based hierarchy prevalent in the high society of the Vienna in which he was raised, and he appears to have seen the emancipation of the working classes as a sign of cultural decline (LC, pp. 7-11; CV, pp. 8f, 81).

Some of these criticisms, however, are perhaps somewhat unfair. Whilst Wittgenstein does talk of certain traits as typically Jewish, he does not give an explicitly negative evaluation of these traits (Schroeder, S. 2006a, p. 11 fn2). Against this it has been argued that Wittgenstein uses his remarks on Jewishness as a device for self-criticism (Monk 1991, p. 316) making it at least questionable that these remarks were entirely neutral observations (Szabados 1999, p. 4). In mitigation, however, it ought to be noted that even if these remarks were covertly derogatory, they are largely confined to a short period in 1930-1, before the rise of Nazism, after which time such remarks took on a graver character in the common psyche.[5]

In a number of places, Wittgenstein distances himself from any so-called 'misogyny' (Drury 1981, p. 91; Monk 1991, p. 313). Moreover, in the context of a time before the rise of feminism and women's liberation, to label a less egalitarian attitude than is prevalent today 'misogynistic' would seem to be harsh at best. In fact, even if some of Wittgenstein's remarks on gender are questionable by modern standards, and might, therefore, be labelled 'sexist', to call them 'misogynistic' would be a misuse of the word, inasmuch as his remarks certainly do not express *hatred* towards women.

Whilst Wittgenstein's first-order moral and political musings have had (some would say mercifully) little impact on subsequent moral thought, he nonetheless had a not

5 It has been argued that the fact that Wittgenstein ceased to make such remarks after this time indicates that part of the value of Wittgenstein's philosophy lies in its having helped him to shed his moral prejudices along with his philosophical ones, and thus in its being apt for helping us to do the same (Szabados 2010). It seems on the whole more likely, however, that his ceasing to make such remarks was a function of the historical context. Moreover, as I argue in Chapter One, freedom from conceptual confusion does not necessitate moral 'improvement', at least inasmuch as this is taken to mean that it necessitates any particular moral conclusions, although it *can* help each agent to judge in conformity with his or her own moral standards.

inconsiderable, yet oft-overlooked influence on some major contributors to normative ethics. Of particular note in this regard are G.H. von Wright, Elizabeth Anscombe and Philippa Foot. Whilst all of these writers were well-versed in Wittgenstein's work (indeed, as Wittgenstein's literary executors, von Wright and Anscombe were undoubtedly authorities on his work), they were influenced not so much by his overall conception of the nature and role of philosophy as by his general philosophical style and personality, and by specific aspects of his philosophy, in particular his explorations of psychological concepts, and his anthropological approach to language, which they extrapolated to their treatments of normative ethics (Glock 2015, p. 106).

By contrast with these Wittgenstein-inspired moralists, Wittgenstein himself did not merely remain silent on normative ethics, but appears to have been actively hostile towards normative ethical theorizing. He seems to have thought that many of the alleged problems of moral philosophy were bogus (Rhees 1965, p. 21) either because the examples considered are so thinly described and removed from everyday experience as to make judging their moral implications impossible, or else because they arise from attempts to 'dig below bedrock' in trying somehow to justify basic moral principles (Waismann 1965, p. 15; cf. Pleasants 2009). This isn't, however, to say that he thought it unimportant to think carefully about moral problems, or that moral questions could never be answered in the abstract (Wittgenstein was more than happy to discuss hypothetical moral problems (Rhees 1965, p. 22)), thus his hostility towards philosophical ethics should not be taken as a kind of irrationalism, but more as an expression of exasperation with the way (and, perhaps, the spirit (cf. Glock 2015, p. 106)) in which most moral philosophers of his day conducted their investigations.

More importantly, though, Wittgenstein's distaste for normative ethical theorizing seems to have stemmed from his metaethical convictions. Wittgenstein, as I shall argue,

took a broadly anti-realist stance on ethics. Many different ethical outlooks are possible, and these diverse outlooks cannot be measured for 'correctness' against an independent moral reality. Hence, there can be no such thing as an ethical 'theory' in any scientific sense, nor in the action-guiding sense that many moral philosophers have aspired to.

Consequently, Wittgenstein's ethical interests were concentrated mainly on metaethics. This focus is a result not only of Wittgenstein's specifically ethical views, but also of his conception of philosophy in general. On a Wittgensteinian account, philosophy is a matter of conceptual clarification. Whilst, then, from a Wittgensteinian perspective, one might well seek to clear up specific conceptual confusions which are a hindrance to first-order ethical thought, this perspective offers nothing particularly unique in the way of specifically moral argument. For, as I shall argue, genuine moral problems are not, like other philosophical problems, conceptual confusions, which can be dissolved by reference to an objective standard of correctness, viz. grammatical rules. Instead, the standard of correctness for moral judgements is subjective, having to do not with facts about the English (or any other) language, but with our feelings and attitudes towards actual and hypothetical actions, forms of behaviour and character traits. Thus, whilst there is a definite affinity between moral reflection and conceptual clarification, the former is not simply a species of the latter's genus.

Wittgenstein's influence on metaethics, whilst perhaps more resounding than that on normative ethics, is almost equally overlooked. Perhaps his most important legacy in this regard is his influence on the erstwhile non-cognitivist orthodoxy of the middle part of the twentieth century.

This orthodoxy is usually credited as having begun, in its modern form, in 1923, one year after publication of *Tractatus Logico-Philosophicus*, with the publication of C.K. Ogden and I.A. Richards' *The Meaning of Meaning* (Ogden & Richards 1923). Ogden was

responsible for publication of the *Tractatus*, and *The Meaning of Meaning* aimed, in part, to build on Wittgenstein's insights therein, whilst rejecting his mysticism and his 'dissatisfaction with ordinary language' (Ogden & Richards 1923, p. 255).[6]

The orthodoxy continued via (amongst others) A.J. Ayer's *Language, Truth and Logic* (1936). Although it is often said that Ayer was indirectly influenced by Wittgenstein through the logical positivism of the Vienna Circle, which was itself in no small part influenced by the *Tractatus* (Russell 1959, p. 160), Ayer, in fact, acknowledges in the very first sentence of the preface to *Language, Truth and Logic*, that the ideas in the book are "derived from the doctrines of Bertrand Russell and Wittgenstein" (Ayer 1926, p. 9).

Following Ayer came Charles Stevenson's 'The Emotive Meaning of Ethical Terms' (1937), and later *Ethics and Language* (1944). The influence on Stevenson was no less direct, Stevenson having studied under Wittgenstein at Cambridge in the 1930s, and, indeed, having been drawn to philosophy from his initial interest in English literature in part by Wittgenstein's influence (Boisvert 2015).

The orthodoxy perhaps culminated in R.M. Hare's *The Language of Morals* (1952) and *Freedom and Reason* (1965). Hare was also much indebted to Wittgenstein for his linguistic approach to philosophy.

Indeed, whilst non-cognitivism's hegemony may now have come to an end, Wittgenstein's influence remains strong with its contemporary exponents (see esp. Blackburn 1990 & 1998).

This influence, however, does not pull uniformly in one direction. At non-cognitivism's zenith, Hare's prescriptivism was already coming under attack from

6 In the year of its publication, Wittgenstein read and disapproved of Ogden and Richards' book (Baker & Hacker 1985, p. 170; CL, p. 137). He did not, however voice dissatisfaction specifically with their non-cognitivism. His dissatisfaction was perhaps most likely with their general theory of meaning, according to which words have the meanings they do on account of their causal role in eliciting certain behavioural responses (Ogden & Richards 1923, ch. 3). Although one can only speculate, this is, perhaps, what Wittgenstein was referring to when he wrote to Russell of the book that "philosophy...is not as easy as that!" (CL, p. 137), i.e. philosophy cannot be reduced to natural science.

Wittgensteinian quarters, perhaps most notably from Foot.[7] More often noted, however, Peter Geach, also a student of Wittgenstein's, famously used ideas he inherited from Wittgenstein's mentor, Frege, in constructing the infamous 'embedding problem', thus becoming arguably non-cognitivism's most influential critic (see Geach 1958 & 1965), and John Searle, who arrived independently at a similar criticism of non-cognitivism (Searle 1969, pp. 150-3), was also heavily influenced by Wittgenstein.

Whilst the problems for non-cognitivism raised by Geach and Searle are serious, they make no direct claim to being based on Wittgenstein's philosophy.[8] Hence, in the rarefied domain of Wittgensteinian ethics, an attack on non-cognitivism from a different quarter usually takes centre stage.

Chief amongst the advocates of this Wittgenstein-inspired assault on the former non-cognitivist stronghold are Sabina Lovibond and John McDowell,[9] whose self-proclaimed 'anti-non-cognitivism' is explicitly based upon ideas drawn from Wittgenstein. If there is presently anything approaching an orthodoxy in Wittgensteinian ethics, it is a school of thought stemming from the trunk of Lovibond and McDowell's critique of non-cognitivism.

Apart from his influence in the debate between non-cognitivists and 'anti-non-cognitivists', Wittgenstein has also spawned a school of 'conceptual relativists' in metaethics, whose most notable adherents are D.Z. Phillips & H.O. Mounce (1969) and Robert Arrington (1989), and he has also been cited in support of 'moral particularism', by, most notably, John McDowell (1979, 1981) and Jonathan Dancy (1993, p. 83; 2004, p.

7 See Chapter Seven for a discussion of one of Foot's main criticisms of non-cognitivism.

8 Glock (2015, p. 111) claims that the embedding problem can be seen as an application of Wittgenstein's pleas to consider the overall use of a term in our discourse, rather than simply isolated applications. Whilst I will argue something similar, any claim that this was what the progenitors of the problem had in mind is rendered rather implausible by the fact that that Searle (2007, p. 10), at least, explicitly denies that Wittgenstein was aware of such a distinction (see Chapter Four for further discussion of this point).

9 See esp. McDowell (1985, 1990) and Lovibond (1983). McDowell and Lovibond are backed up perhaps most notably by (amongst many others) David Wiggins (1987).

197). [10]

From the profusion of Wittgenstein-inspired approaches to metaethics, often reaching seemingly diametrically opposed conclusions, it seems clear that the nature of Wittgenstein scholarship is to abhor a vacuum. In what follows, then, my intention is to attempt to add a little more air to the void left by Wittgenstein's silence on ethics. In doing so, I hope not only to make it possible for my own voice to be heard, but also, if possible, to allow Wittgenstein's own thoughts on the subject to resonate a little louder than hitherto they have.

In this thesis, my interest, like Wittgenstein's, will be focused on metaethics. In summary, I will argue that Wittgenstein's thoughts, both about ethics specifically and about language and philosophy in general, support a fairly simple-minded kind of subjectivism. The major impediments to acceptance of such a naive account of moral terms, I shall argue, stem from misunderstandings about concepts such as (amongst others) meaning and truth and the relationship between the two, and about the role of analyses in philosophy and the standards according to which they should be judged.

The task is not simply one of speculation regarding what Wittgenstein might have said on the topic, had he devoted more attention to the matter. Whilst this thesis is in part exegetical, any claim to be able to extract or extrapolate a coherent metaethic from Wittgenstein's remarks would be both hubristic and, ultimately, futile, given the dearth of material from which to begin.

Neither is the task to take on every facet of the various Wittgenstein-inspired metaethics. Whilst I make some criticisms of extant Wittgensteinian views and draw on

10 It is beyond the scope of this thesis to give a thorough treatment of the debate over moral particularism. It is not a position which Wittgenstein ever expressed himself, but one which has been inspired by his thoughts, particularly on 'rule-following'. Suffice it to say that it seems to me that the particularist's 'holism of reasons' thesis is broadly correct, but it is unclear that full-blown particularism (the claim that there are no moral principles, or that reliance upon moral principles leads somehow to moral error (Dancy 2013)) follows from it. In any case, particularism appears at odds with much of what Wittgenstein had to say about moral judgements being non-contingent expressions of rules (see Chapter Two).

some of the insights of Wittgenstein's diverse group of followers, where this serves to strengthen my own argument, the task is neither the negative one of debunking every alternative view, nor that of choosing one to defend. Rather, what follows is an attempt, guided largely by Wittgenstein's own reflections on ethics, as well as his thoughts on other areas of philosophy, and, indeed, on the nature and role of philosophy in general, to breathe new life into an account of ethics not generally much associated with Wittgenstein and often thought hopelessly outdated, flawed and naive.

Metaethics in general, though, is sometimes thought to be at best of secondary import, in comparison with normative or applied ethics, or, at worst, an unwelcome distraction from the more obviously important task of resolving genuine moral problems, and the charge seems to have particular potency in relation to a Wittgensteinian approach. For, since it isn't immediately clear how 'trivial' grammatical reminders could be of any great importance or interest, it has often been argued that Wittgenstein's conception of philosophy impoverishes the subject as a whole to the point of bankruptcy (See e.g. Russell 1959, p. 126; Popper 1952; Marceuse 1964, pp. 177f).

In Chapter One, then, I attempt to motivate the task of outlining a Wittgensteinian metaethic, by exploring what value such a project might have. I consider a prominent interpretation of what, on a Wittgensteinian account, gives philosophy value, namely its role as a 'therapy' for philosophical 'torment'. I argue that, whilst, *pace* some commentators (see e.g. Bouwsma (1961)), such an aim ought not, on a Wittgensteinian account, to be taken as *definitive* of philosophy, philosophy's ability to achieve this aim can, indeed, give it value, although this shouldn't be taken to be philosophy's *only* source of value.

Thus, I also consider the claim that there is, for all thinking people, an intrinsic value to truth. I argue that, contrary to popular opinion, its ability to lead us to the truth *can* explain Wittgensteinian philosophy's value to some extent, although, once again, to take

this as philosophy's *only* source of value might leave the subject open to the time-honoured charge that it is a rather esoteric pursuit.

I therefore move on to examine Wittgenstein's comments about the possibility of conceptual investigations providing a 'foundation' of science and mathematics. I argue that there is a sense in which philosophy can set these areas on a surer footing, although it is not the sense in which philosophers have traditionally taken it to do so.

I end by applying the notion that conceptual clarification may play a 'foundational' role (in this attenuated sense) to ethics. I argue that, contrary to the charges often levelled against it, metaethics-as-conceptual-clarification *can* be of great importance to substantive ethical thinking, since conceptual confusions involving moral concepts *can* impact upon one's substantive moral outlook. This is not to say that conceptual clarity will *necessarily* have an impact on a person's ethical outlook, let alone a morally positive impact. It *can*, however, help each person to judge in conformity with his or her[11] own moral standards. Thus clarity about moral concepts is important for individuals who seek to live morally.

Whilst, in the first chapter, I argue that we ought not, primarily, to be interested in what Wittgenstein had to say specifically about ethics, but in the application of his conception of philosophy to this area, it would nonetheless be remiss to ignore what little he did have to say on the subject. In Chapter Two, then, I examine Wittgenstein's remarks on ethics chronologically, in order both to see how his ethical thought developed along with his changing views on the nature and role of philosophy as a whole, and to narrow down the metaethical options, thus guiding the argument towards a starting position.

I argue that one of Wittgenstein's major concerns in the *Notebooks* and the *Tractatus* lay in reconciling his 'internalism' with the mind-independence of ethics. This is, in essence, the modal problem for simple subjectivism, which later became a founding

11 From here on in, for the sake of brevity, I shall use masculine pronouns only. This is merely a stylistic preference and is not intended as any kind of rebuke to those who choose to use only feminine pronouns.

motivation for non-cognitivism. Wittgenstein's solution to the problem at the time, I shall argue, was to place the attitudes upon which ethics is founded 'outside the world', in the ineffable sphere of the metaphysical. This concern, I argue, continues through to the 'Lecture on Ethics'.

The focus of later writings, however, is somewhat different. In his discussion of ethics in his Cambridge lectures, some four years after the *Lecture on Ethics* was penned, Wittgenstein approaches the subject mainly at a tangent, via discussions of aesthetics, religion and ritual practices, an approach which, from here on in, becomes a *Leitmotif*.

It has been argued that, during this period, Wittgenstein gave up his adherence to the fact/value distinction in favour of a kind of contextualist variant of naturalism (Stern 2013). I argue, however, that there is reason to doubt this exegetical claim, and that, even if Wittgenstein did briefly adopt such a position, he was mistaken to do so, since such an approach is incompatible with his later views in other areas. I argue that, whilst it is indeed a consequence of Wittgenstein's later view of language that, in a particular context, a moral term can take on a descriptive meaning, due to the fact that speakers often share, or at least know, one another's values, this does not make a token use of the term on such an occasion a *mere* description, nor does the fact that moral judgements can have such a descriptive function overturn the distinction between fact and value.

I go on to discuss Wittgenstein's lectures on aesthetics, in which it may seem that he took the key to understanding aesthetics to lie in turning away from 'thin' concepts, such as 'beautiful', and, in ethics, 'right', 'wrong', 'good' and 'bad', as some of his followers have suggested (see e.g. McDowell 1979, 1990). I argue, however, that Wittgenstein's change of emphasis at this point was due more to his ceasing to see such 'thin' concepts as particularly problematic, having already settled upon a broadly expressivist treatment. Wittgenstein was, at this point, more concerned with 'genuine' aesthetic (and moral)

controversies arising between those who shared a 'cultured taste' than with the simple clashes of attitudes that occur when one person thinks a thing beautiful and another ugly.

Lastly, I consider Wittgenstein's most mature views. I argue that, whilst Wittgenstein took (some) moral judgements to be expressions of rules, he did not, as some commentators (Phillips & Mounce 1969; Arrington 1989) have thought, subsume moral rules under grammatical rules, and did not, as these commentators do, espouse a form of relativism, but in fact explicitly rejected such an approach.

I conclude that, whilst Wittgenstein never had a fully worked out position on metaethics, the direction of his thought on the matter remained relatively constant throughout his life, and that, despite realist animadversions, Wittgenstein's own direct contributions to the topic are broadly anti-realist or expressivist.

In Chapter Three, I go on to discuss the 'expressivism' for which Wittgenstein is perhaps best known, namely that in relation to first-person present-tense psychological utterances.

I outline Wittgenstein's objections to the 'traditional conception of the mental' as a realm of logically private 'inner objects', and his counter-suggestion, that psychological terms are generally used to replace and extend natural expressive behaviour.

I question a recent interpretation of Wittgenstein's lesser-known comments about 'descriptions' of psychological states, arguing that the point of Wittgenstein's discussion in this regard is not to endorse a 'liberal' view of 'description', consonant with our everyday use of the term, but rather to highlight that the variegated nature of our everyday concept of 'description' can be philosophically misleading, even to the extent that we may be better off avoiding the term in the context of a philosophical discussion about psychological utterances.

There are, however, cases in which Wittgenstein grants that there can be

expressions of pain or fear which are 'genuine' descriptions, inasmuch as they allow for the possibility of doubt and error, and these, I suggest, may constitute a model for our moral vocabulary.

In Chapter Four, I discuss attempts to place one of subjectivism's major obstacles – the modal problem, concerns similar to which so occupied the earlier Wittgenstein – back in the path of the non-cognitivism which it spawned. I argue, following Mark Schroeder (2010b & 2014), that such attempts fail, since they fail to take into account a central tenet of contemporary expressivism: the 'Parity Thesis'.

This discussion, though, reveals the importance to contemporary non-cognitivism of giving an account of 'Moore's Paradox'. This was a subject which Wittgenstein felt to be of tremendous philosophical importance, and which he consequently explored in some depth. I therefore outline Wittgenstein's treatment of Moore's Paradox, arguing that at its crux lies a distinction between the meaning of an utterance and the meaning of a sentence.

I conclude that, whilst non-cognitivists may have given a correct account of the meaning of a simple moral utterance, a response to the Frege-Geach point cannot follow from it.

I begin Chapter Five by summarising some of the main responses to the Frege-Geach point. I argue that these fail, since most capture the use of moral terms incorrectly, and even the best amongst them captures at most a part of their use. Thus the non-cognitivist still has work to do in responding to the point.

I then argue that a kind of 'idealizing subjectivism' fares better in this regard. Such views, however, are subject to serious challenges of their own – challenges which, as I argued in Chapter Two, concerned the earlier Wittgenstein. After discussing some more peripheral objections, then, I turn to the task of tackling the disagreement and modal problems.

I argue that the 'disagreement problem' can be seen off by Wittgenstein's minimalism about truth. In arguing this, I contrast my position with the 'anti-anti-realist' view that Wittgenstein's truth-minimalism leads to a kind of across-the-board realism. Wittgenstein's comments on truth, I argue, were intended to make precisely the opposite point.

I then argue that, with the distinction between sentence and utterance meaning in hand, we can take the modal problem not as an objection to a philosophical theory, but as a comment on the grammar of moral sentences. Whilst the use of a moral sentence is, in most respects, like that of a psychological self-description, it differs in some respects, most notably that when such a psychological sentence is transformed into the subjunctive, it can be used to talk about a hypothetical difference either in one's attitudes or in the properties of the action (etc.) under consideration, whereas the corresponding transformation of a moral sentence may only be used to talk about a hypothetical difference in the features of the action under consideration, and not in one's attitudes. Such a caveat, I claim, lies at the very heart of Wittgenstein's later philosophy, in his reminder that 'meaning is use', and the addition of a similar caveat to our analysis of moral terms allows us to retain a subjectivist analysis, without having to deny the point made by the modal problem.

My solution to the modal problem, however, may seem an arbitrary fix. Indeed, David Enoch (2005) has argued that all 'idealizing subjectivist' metaethics suffer from a similar flaw. Thus in Chapter Six, I consider Enoch's objection, and its application to the appearance of an *ad hoc* fix in my argument.

I argue that much of the force of Enoch's argument is dissipated by the recognition that it contains an equivocation between two distinct senses of 'a rationale for idealization': a rationale for analysing a concept in a particular way, and a rationale for engaging with moral concepts as the 'idealizing subjectivist' claims them to be.

I argue that the 'idealizing subjectivist' needs no more rationale for the use of the idealization manoeuvre in his analysis of moral concepts than that his analysis is both plausibly correct and useful in resolving a particular philosophical problem and that, *pace* Enoch, the subjectivist's claim to have produced a correct analysis *can* be justified by appeal to our practice with moral terms. This argument applies equally to my contention that we may restrict the application of an analysis as suggested in the previous chapter.

Moreover, there is, on an idealizing subjectivist analysis, a compelling rationale for engaging with moral concepts, since doing so allows one more fully to fulfil one's desires, or more easily to live according to one's fundamental attitudes. Hints towards this latter kind of rationale, I shall argue, can be found in David Sobel's (2009) response to Enoch. Sobel, however, like Enoch, fails to distinguish the two kinds of rationale and thus ultimately fails to provide a rationale of either kind.

In Chapter Seven, I discuss the further objection, given prominence by Philippa Foot, that a metaethic such as the one I outline is too 'liberal' regarding what it allows to fall under the concept of a moral judgement. I argue that Wittgenstein's discussions of ritual practices and aesthetics constitute a double-edged sword in this regard, for they suggest that, whilst, *pace* Foot, there are no strict criteria according to which an attitude may be called a moral one, there are, nonetheless, more conceptual constraints on legitimate moral judgement than Foot's opponents allowed. Thus, there is no conceptual room for Foot's utterly idiosyncratic hand-clasper, but room can be made with the addition of only a minimally described context.

I conclude that, using Wittgenstein's later thoughts on language to address his earlier puzzlements about ethical concepts, what emerges is a plausible version of subjectivism, which accounts for the action-guidingness of moral judgements, whilst not being objectionably mind-dependent, *ad hoc* or 'liberal'.

I

Hunting for Food: A Wittgensteinian Perspective on the Importance of Metaethics

Developing a Wittgensteinian account of ethics is a project that, for a number of reasons, seems to stand in need of special justification. Firstly, Wittgenstein himself wrote very little on the subject of ethics, and what he did write is confined largely to the period before his new conception of philosophy had been fully worked out. As I shall argue in this first chapter, however, we ought not to be interested primarily in exegesis of Wittgenstein's comments specifically on ethics, but in the application of his philosophical method to confusions involving moral concepts. Thus there is scope for a Wittgensteinian metaethics.

Even metaethics as traditionally conceived, however, is seen by some as at best of secondary importance when compared with the task of providing answers to substantive ethical questions, and such a worry seems especially acute in the case of a Wittgensteinian account. For it is commonly held that Wittgenstein's brand of 'philosophy-as-conceptual-clarification' lacks the depth and seriousness of genuine metaphysical enquiry. These concerns are heightened by the fact that, contrary to popular belief, Wittgenstein had a high opinion of much traditional philosophy. That Wittgenstein should take such a positive attitude towards what he repeatedly branded as 'nonsense', and that, having such an attitude towards traditional philosophy, Wittgenstein should nonetheless see fit to set about debunking it, however, seems nothing short of paradoxical.

My aim in this chapter, then, is to dissolve this seeming paradox and to suggest how philosophy in general, and metaethics in particular, can be of value on a Wittgensteinian account.

A Method Has Been Found

Any attempt to put forward a Wittgensteinian approach to ethics seems in very real danger of being bereft of a subject matter. Wittgenstein wrote almost nothing which appears directly relevant to the corpus of moral philosophy. In fact, both the frequency and extent of Wittgenstein's explicit treatments of ethics suffered a notable decline throughout his career, leaving the scholar with the distinct impression that he became less interested in ethics as his ideas about the nature of philosophy evolved.

By far his most sustained treatment of ethics is to be found in his wartime notebooks (NB, pp. 76-91). In Wittgenstein's first great work, *Tractatus Logico-Philosophicus*, ethics is discussed only briefly, as one amongst a number of philosophical issues which he takes to have been shown by the argument of the rest of the book to be 'ineffable' (TLP 6.4 – 6.45). In fact, Wittgenstein's only other substantial contribution to moral philosophy and, indeed, his only work devoted solely to consideration of ethical questions, comes in the form of his 'Lecture on Ethics' (LE), which was delivered some fourteen years after his closing remarks on ethics were entered into his notebooks[12] and consists of only nine pages of text. Of the nearly seven hundred remarks contained in his *magnum opus*, *Philosophical Investigations*, there is but one solitary remark in which direct comment is passed on ethics (PI §77)[13] and this is typically cryptic and laconic, and appears in the midst of a discussion regarding a seemingly unrelated subject.

Given Wittgenstein's reticence on matters ethical, the question arises why we should be much interested in his thoughts on the matter. The answer is that we ought, in fact, not to be interested primarily in what he himself said on the matter. Indeed,

12 Wittgenstein's final remark on ethics in the Notebooks was penned on 10th January 1917 and the 'Lecture on Ethics' was composed some time between September 1929 and December 1930.

13 There is at least one other reference to ethical matters in PI §304. This remark, however, is not concerned primarily with ethics, but merely uses ethics as an example to illustrate a point about the multifarious functions of language.

Wittgenstein thought that anything he might have to say in response to any specific philosophical problem was, at best, of secondary importance. In his notes on Wittgenstein's Cambridge lectures (one of the few other places where Wittgenstein expresses any opinion about moral philosophy (Moore 1955, pp. 16-19)) G.E. Moore recalls that

> As regards his own work, he said it did not matter whether his results were true or not: what mattered was that "a method had been found". (Moore 1955, p. 26)

There is general consensus amongst Wittgenstein scholars that the central import of Wittgenstein's philosophy lies in his novel and subversive conception of philosophy's aims and methods.[14] Wittgenstein's most important contribution to philosophy is arguably to have highlighted the need for philosophers to shift the emphasis of their enquiries from questions of truth to questions of sense (CV, p. 3). Philosophy is no rival to the natural sciences in discovering any kind of fact about the world. It can neither justify nor undermine our empirical knowledge, and it cannot give us access to a realm of metaphysical facts, discoverable *a priori*. Instead, philosophy's proper role lies in clarifying our concepts, which, in the context of philosophical debates, we are tempted to misuse, leading us into subtle forms of unintelligibility.[15]

This conception of philosophy would appear to narrow the scope of Wittgensteinian ethical enquiry even further. For it would appear to rule out its having anything of use to contribute to first-order ethical questions, since (if the argument of this thesis is correct)

14 There is, of course, by no means complete agreement as to what Wittgenstein's conception of philosophy amounted to. This, however, is not the place to embark upon an in-depth defence of any particular reading of Wittgenstein, so I shall make certain assumptions about how Wittgenstein is to be read.

15 That many philosophical pronouncements were neither true, nor false, but nonsensical was a *leitmotif* of both his earlier and later works (cf. Glock 2004, p.223; cf. PI §464 – "My aim is: to teach you to pass from a piece of disguised nonsense to something that is patent nonsense").

such questions are not conceptual questions about the meanings of ethical terms.

There is, however, scope for the application of Wittgenstein's methods to the problems of metaethics, which deals not with substantive moral claims, but with questions such as 'What is the nature of a moral judgement?' and 'Are there moral truths and if so, how can we come to know them?'. On Wittgenstein's account of philosophy, such questions are conceptual questions about the meanings of moral terms. Whilst exegesis of Wittgenstein's ethics may be of limited interest, then, there is still scope for the development of a Wittgensteinian metaethics.

The Seeming Poverty of Conceptual Investigations

A worry may yet remain, however, as to why we should be much interested in pursuing such a project. For, after its heyday in the early part of the twentieth century, metaethics increasingly came to be seen as of lesser importance than normative and applied ethics, or even as an unwelcome distraction from the important business of providing answers to genuine moral questions, or developing moral theories.[16]

A Wittgensteinian approach to metaethics seems peculiarly vulnerable to this charge. For Wittgenstein's brand of philosophy-as-conceptual-clarification has struck many as uniquely impoverished. This feeling was perhaps most plainly expressed by Wittgenstein's mentor, Bertrand Russell, who said of Wittgenstein's later philosophy:

16 This is not a view that is very often given voice explicitly in print, but more a general feeling that appears to enter into the philosophical Zeitgeist from time to time, arguably, for example, in the move away from metaethics towards applied ethics in the 1970s (cf. Delapp 2016).

> I have not found in Wittgenstein's *Philosophical Investigations* anything that seemed to me interesting and I do not understand why a whole school finds important wisdom in its pages. Psychologically this is surprising. The earlier Wittgenstein, whom I knew intimately, was a man addicted to passionately intense thinking, profoundly aware of difficult problems of which I, like him, felt the importance, and possessed (or at least so I thought) of true philosophical genius. The later Wittgenstein, on the contrary, seems to have grown tired of serious thinking and to have invented a doctrine which would make such an activity unnecessary. I do not for one moment believe that the doctrine which has these lazy consequences is true. (Russell 1959, p. 126)

An aside is warranted at this juncture in order to point out a *non sequitur* which is common to much of the criticism levelled against Wittgenstein's philosophy. Russell inveighed against Wittgenstein essentially on the grounds that his philosophy, if correct, entails that the philosopher can be no scientist. Karl Popper (1952) took issue with Wittgenstein on similar grounds (disagreement on this matter being at least partly responsible for an infamous altercation involving a fire iron.)[17] In a similar vein, Marcuse (1964, pp. 177f) rails against Wittgenstein on the grounds that his philosophy would entail that the philosopher cannot be an agitator for social change through the introduction of subversive new concepts.[18]

But the fact that a philosophy has consequences which one finds unpalatable has no bearing whatsoever on its correctness. Russell himself, when asked what wisdom he should like to impart to future generations said:

17 See Edmonds & Eidinow 2002 for an entertaining and informative account of this incident.

18 It is not at all clear that Marcuse was *right* about this being a consequence of Wittgenstein's philosophy, but even assuming for the sake of argument that he *was* right, no conclusion follows about the correctness of Wittgenstein's philosophical method.

> When you are studying any matter, or considering any philosophy, ask yourself only what are the facts and what is the truth that the facts bear out. Never let yourself be diverted either by what you wish to believe, or by what you think would have beneficent social effects if it were believed, but look only, and solely, at what are the facts. (BBC interview – See link in footnote[19])

It seems, then, that in his criticism of Wittgenstein, Russell failed spectacularly to follow his own advice.

Whilst these criticisms fail to address the logical *correctness* of Wittgenstein's philosophy, they might, nonetheless, appear to bring into question its *value*. For one might agree that Wittgenstein has come up with the correct philosophical method, yet question the point, on a Wittgensteinian account, of continuing to philosophize.

In the *Investigations*, Wittgenstein anticipated such questions regarding the value of his radical new view of the nature of philosophy:

> Where does our investigation get its importance from, since it seems only to destroy everything interesting, that is, all that is great and important? (As it were all the buildings, leaving behind only bits of stone and rubble.) (PI §118)

His answer at this point, however, is dismissive:

> What we are destroying is nothing but houses of cards and we are clearing up the ground of language on which they stand. (PI §118)

19 https://www.youtube.com/watch?v=ihaB8AFOhZo

This glib response may seem initially attractive. After all, if traditional philosophy is nonsense, then it can hardly have the value that philosophers have traditionally imputed to it. However, it still remains to be shown what value there is in pursuing Wittgenstein's alternative.

Moreover, Wittgenstein was not always so dismissive of philosophical 'houses of cards'. He was not, as is sometimes supposed, utterly contemptuous of traditional philosophy. In conversation with Maurice Drury in 1930, Wittgenstein remarked

> Don't think I despise metaphysics. I regard some of the great philosophical systems of the past as amongst the noblest productions of the human mind. (Drury 1984, p. 105)

In fact, Wittgenstein made repeated reference to the 'greatness' of some traditional philosophers (Moore 1955, p. 26; Drury 1984, pp. 105f).

If, however, past philosophical systems are, as Wittgenstein suggests, nonsensical, then this respect is puzzling in itself. For it is not immediately clear how nonsense could be important and worthy of respect. Moreover, if nonsense can be important and worthy of respect, this serves only to intensify the pressure to justify employing Wittgenstein's method in debunking it.

What, then, are we to make of such seemingly paradoxical remarks? Wittgenstein's respect for traditional philosophy certainly wasn't, for example, based upon the intellectual difficulty involved in its pursuit. He distinguished firmly between 'great philosophers' and 'very clever people' (Drury 1984, p. 106). One thing that, to Wittgenstein's mind, made a

philosopher 'great', it seems, was that his philosophy was an attempt to grapple with a genuine problem of life, or, perhaps, a confused attempt to express something which is of genuinely profound significance (cf. Johnson 1989, p. 23). Thus he praised William James as a great philosopher, not on the ground that he was clever, but on the ground that "he was a real human being" (Drury 1984, p. 106). Indeed, he praised thinkers whose work had such a practical, or ethical, dimension even where he thought their arguments ridiculous. Thus he said of Lenin, for example, that

> ...[his] writings about philosophy are of course absurd, but at least he did want to get something done. (Drury 1984, p. 126)

Wittgenstein appears also to have held that traditional philosophy could be a powerful psychological remedy in much the same way that he held religion could be. The above quotation praising metaphysical systems is taken from a wider conversation on the subject of a chapter from Schopenhauer entitled 'On Man's Need of Metaphysics' (Schopenhauer 1818, pp. 359-395) and is immediately preceded by the comment that

> I think I can see very well what Schopenhauer got out of his philosophy. (Drury 1984, p. 105)

It is not hard to see how someone as transfixed by the pain and suffering of life as was Schopenhauer could find some cold comfort in the notion that this is not an inexplicable brute fact of human existence, but the work of a maleficent 'World-Will'

(Schroeder, S. 2007, p. 450).[20]

It was on similar grounds that Wittgenstein was prepared to let religious metaphysical beliefs go unchallenged, so long as they were not taken to be supported by evidence or reason (LC, p. 58; Schroeder, S. 2007; cf. Drury 1984, p. 102). Indeed, Wittgenstein elsewhere suggests that philosophical metaphysical beliefs, like ritualistic beliefs, should neither be defended nor derided, so long as they do not become 'superstitious', in being taken pseudo-scientifically, to entail truths about the course of our experience (RFGB, p. 116).

The tension between this respect for metaphysics and Wittgenstein's insistence on debunking it, however, cries out for resolution, and the tension is only heightened by the fact that Wittgenstein's own philosophy may appear to mark him out squarely, not as a 'great' philosopher, but only as a 'very clever man'. Indeed, in a lecture recorded by G.E. Moore, Wittgenstein himself once said that

> ...it was now possible for the first time that there should be “skilful” philosophers, though of course there had in the past been “great” philosophers.

and that

> ...philosophy had now been “reduced to a matter of skill”... (Moore 1955, p. 26)

20 Indeed, reading Wittgenstein's wartime notebooks and diaries gives one the distinct impression that his ethics and metaphysics at that time were intimately connected to his deep psychological need to feel that the good life was not dependent either on the harsh contingencies of the world, or on his own ability to do anything about these. Thus, Wittgenstein's ethics in the *Tractatus* appeared to him to offer a way to be happy in spite of the awful situation in which he found himself (Schroeder, S. 2006a, pp, 102f).

For it is not immediately obvious that Wittgenstein's philosophy offers anything that could have the kind of worldly import, the quest for which he so respected in traditional philosophy. Thus it may seem that his philosophy is just the kind of sterile word-play for which he had no time.

The 'Therapeutic Reading'

In response to such worries, many scholars have seized upon Wittgenstein's comments about philosophy as a kind of 'therapy'. In a particularly intense section of the *Investigations*, in which he sets out the purpose of his new way of doing philosophy, Wittgenstein says

> The results of philosophy are the uncovering of one or another piece of plain nonsense and of bumps that the understanding has got by running its head up against the limits of language. These bumps make us see the value of the discovery. (PI §119)

Such passages have inspired a burgeoning school of thought, according to which the aim of philosophy is 'therapeutic'. As Gordon Baker (in his latter days, a proponent of a 'therapeutic reading') has it:

> ..dealing with compulsions, obsessions, prejudices, torments,... is the *proper business* of philosophy. (Baker 2004, p. 219)

Wittgenstein often likened the problems of traditional philosophy to a 'disease' (PI §§255, 593). The pronouncements of traditional philosophy, Wittgenstein declared, were not (by the lights of his new method) philosophy at all, but "something for philosophical *treatment*" (PI §254, Wittgenstein's emphasis). He described the aim of philosophy as the alleviation of 'mental cramps' (BB, pp. 1, 59, 61), and emphasised the 'torment' felt by those who are racked by philosophical problems (RFM, p. 120, §13; RC, §33; cf. PI §133). Thus, it is argued, philosophy can have value because of the 'therapeutic' effect of removing philosophical confusions.

Indeed, some have gone so far as to claim that Wittgenstein's new conception of philosophy is to be *defined* in terms of the aim of alleviating anguish. (see e.g. Bouwsma 1961) Such a reading is generally based on Wittgenstein's purported claims that there should be no 'theses' in philosophy (PI §128) – that philosophical questions cannot be *answered* (EPB, p. 156; cf. RFM, p. 147), but must be *rejected* (BB, p. 8), and that philosophical problems should not be resolved, but rather *dissolved* (BT 421; Glock 1991, p. 73).

This interpretation of Wittgenstein's conception of philosophy, however, has been roundly criticized by members of what one might call the 'orthodox' school of Wittgensteinian thought, as both exegetically and substantively flawed (Glock 1991; Hacker 2008).

Indeed, the extreme version of the 'therapeutic reading', at least, is inimical to the whole thrust of Wittgenstein's later philosophy as the unveiling of latent nonsense (PI §§464, 524). For it appears committed to the claim that there are no correct or incorrect answers to philosophical problems, but only those which achieve their therapeutic aim and those that do not, leaving it unable to distinguish logically between a good philosophical argument and persuasive rhetoric. (Glock 1991, pp. 83f) The correctness or otherwise of a

statement to the effect that an expression does or does not make sense, though, is determined by the rules of our language. Wittgenstein's continual preoccupation with questions of sense appears strangely irrelevant if the aim of philosophy is not finding correct answers to questions of sense, but merely the alleviation of anguish, which might just as easily be achieved by a course of drugs, or the proverbial 'knock on the head'! (cf. Glock 1991, p. 84). Some arguments, Wittgenstein insists, are to be rejected as just plain wrong (TS 213, p. 408). Moreover, Wittgenstein *does* provide *answers* to some philosophical questions, such as 'What is meaning?' (PI §43) and (correlatively), 'What is understanding?' (PI §§143-84; Glock 1991, p. 74).

Philosophy is no more characterised by the aim of 'treating' the 'anguish' of conceptual confusion than hunting is characterised by the aim of feeding oneself. The aim of hunting is to capture (and kill) an animal. To hunt *is* to act with these aims. The question whether there is any *point* in going hunting is quite different from the question what, logically speaking, the aim of hunting is. If you are not hungry, then going hunting may be a waste of time (although you may yet appreciate hunting as a sport, engaged in for its own sake, or for the thrill of the chase.)

That there are criteria of *correctness* for philosophical pronouncements, laid down by the rules of our language, however, is not inconsistent with there being a separate criterion of *success* for engaging in philosophical activity (cf. Hacker 2008, p. 102). Success, after all, is relative to intentions, and there is no need to deny that the alleviation of philosophical unease can be *a* perfectly respectable aim of practising philosophy.

However, if this is the *only* aim of philosophy, as claim adherents of the less extreme form of the 'therapeutic reading', then it would appear to be rendered a fairly niche pursuit, perhaps only of interest to professional philosophers and not even to most of these, who are not, by and large, *troubled* by philosophical confusion, but, rather, thrive on it.

In the *Investigations*, Wittgenstein famously claimed his aim in philosophy was “To show the fly the way out of the fly-bottle.” (PI §309) Ryle once asked 'What has a fly lost who never got into a fly-bottle?' and it would appear that Wittgenstein's answer would have been 'Nothing' (Schroeder, S. 2007, p. 165). Now, it may be argued that this is immaterial, since we all get stuck in a fly-bottle once in a while. That is, we cannot help but fall into conceptual confusion from time to time (cf. Kenny 2002, pp. 15f). But it does not follow that we all *mind* the fly-bottle that much. Indeed, one might ask 'What has a fly lost who likes it in the fly-bottle?'.

It has been claimed that, even for the philosopher who takes pleasure in his metaphysics, the traditional method of philosophy cannot ultimately be satisfactory (Johnson 1989, p. 22). For, given that traditional answers to philosophical confusions are themselves confused, such answers will simply bring up new problems. In such a case, the philosopher is like a carpet layer, whose carpet has been cut larger than the floor-space and so has a lump in it. He pushes the lump down and is temporarily satisfied that it has gone, until he turns around and the lump has simply reappeared elsewhere in the room. Thus, Wittgenstein says that

> The real discovery is the one that makes me capable of stopping doing philosophy when I want to.—The one that gives philosophy peace, so that it is no longer tormented by questions which bring *itself* in question. (PI §133)

But whilst it may be true that coherence can come only through practising Wittgenstein's method, it is not obvious that people can't be satisfied with their incoherent systems. Indeed, for many, chasing the lumps around the carpet is part of the fun of the game, and does not detract from the psychological comfort that their systems provide.

The therapeutic reading, then, seems at best to make the cult of Wittgenstein one which preaches only to the converted. Wittgensteinian philosophers, however, have generally taken theirs to be a proselytising religion. What, then, justifies reaching out beyond those already seeking redemption?

Truth

To a letter from his sister, in which she had called him a 'great philosopher', Wittgenstein replied, “Call me a truth-seeker, and I will be satisfied” (Monk 1991, p. 3).

Philosophers often appeal to the intrinsic value of truth in order to justify their endeavours. It is sometimes thought that this avenue of response is unavailable to the Wittgensteinian philosopher, since he is not interested in the truth, but merely in clarity. This, however, is not so. Wittgenstein's appeal to philosophers to shift the emphasis of their enquiries to questions of sense is not, as it may seem, an appeal to give up seeking the truth; it is merely an appeal to look for the truth where it can be found using the *a priori* methods of philosophy. The truths in which the philosopher deals are *analytic* truths – truths, that is, which are such in virtue of the meanings (uses) of their constituent words (Schroeder, S. 2009). One answer to the question what a fly has lost who likes it in the fly-bottle, then, is 'The truth!'.

The formulation of analytic truths may not help us to discover any new facts about the world, but it *can* help us to gain an 'overview' of the use of a word, and thus to achieve intellectual clarity.

Wittgenstein valued intellectual clarity exceptionally highly. Indeed, his unwillingness to accept any lack of clarity in his own thinking was a primary reason why

he could not attain the religious faith that he so craved (Schroeder, S. 2007, p. 459). For Wittgenstein, there appears to have been something of an ethical imperative to be clear in one's thinking, seemingly based on his quasi-aesthetic ideals about what constituted good character – an attitude perhaps exemplified in his famous comment that "work on philosophy is…a kind of work on oneself" (BT 161).

In a draft preface to the *Investigations*, Wittgenstein claims that, in contrast to the scientistic thinking of his contemporaries, for whom clarity is at best a means to the end of 'progress', which takes the form of building ever more complex systems, for him clarity is an 'end in itself' (CV, p. 9).[21]

In fact, though, *no* rational person, as such, can be completely satisfied with being conceptually confused. Clarity is, *pace* Wittgenstein's apparent suggestion, an 'end in itself' for *all* thinking people. Inasmuch as the philosopher seeks the truth, then, his incoherent systems cannot ultimately be 'satisfactory', even if he is not actually 'tormented' by this fact; he simply won't have reached his goal.[22] Hence Wittgenstein's philosophy has value inasmuch as it helps us to achieve the end of discovering the truth.

For some, though, this may not be enough. For, that something is an end in itself, does not necessarily make it an overriding consideration. If one accepts that philosophical problems are conceptual confusions, one needn't feel any absolute need to undertake a philosophical investigation. After all, such investigations are difficult and protracted, and if their end result is merely a 'reminder' of how, as a competent speaker, one normally uses words, one might satisfy oneself that such confusions can simply be dismissed as irrelevant

21 The clarity Wittgenstein sought should not be confused with something that is obvious or easy to understand. Wittgenstein's grammatical statements, whilst usually quite precise, do not wear their precision of their sleeve and often (quite intentionally) require a great deal of unpacking. In this respect, Wittgenstein's philosophy can be likened to a mathematical formula, such as Einstein's theory of general relativity – undeniably precise, and beautiful to those who have the ability to see its consequences; completely unfathomable to those who do not.

22 Indeed, Wittgenstein's likening of philosophy to therapy is more plausibly interpreted metaphorically than literally. The sickness of which he spoke needn't be thought of as something that *bothers* a philosopher; it is merely a lack of clarity. (Hacker 2008, pp. 96-99)

to one's everyday life. Where one feels a philosophical difficulty, one might simply rest easy in the knowledge that one's everyday use of the terms involved is in perfect order. If philosophical truths are supposed simply to capture rules of which, as competent speakers of a language, we already have implicit mastery, it is not immediately apparent what *further* importance discovering[23] them might have.

One might, perhaps, argue that some genuinely scientific endeavours are similarly esoteric. The justification for many such pursuits, though, comes not only from the fact that such blue-sky thinking has often been found to pay in the long run, but also from the fact that it tells us *non-trivial* truths, which can be beautiful and awe-inspiring, genuinely giving us a sense of our place in the natural order of things and binding together our knowledge in a way that supports more humdrum scientific discoveries.[24] If this were not so, then it seems likely that the point of such scientific research would appear opaque to many. But it is not immediately obvious exactly how clarity about the meanings of words could share these valuable features of the frontiers of science.

Philosophy, Science and Mathematics

Is there anything more that can be said, then, by way of injecting value into Wittgenstein's philosophy? An answer to this question may, I think, be found in what has been published as Part II of the *Investigations.* Here, Wittgenstein intimates that, in addition to its being an 'end in itself', conceptual clarity might *also* be useful as a means to the end of aiding

23 Whilst some trivial or obvious analytic truths can't be said to be 'discovered' by competent speakers, since not knowing them would betray ignorance of the meaning of (at least some of) the words involved (e.g. 'A vixen is a female fox'), others can indeed be 'discovered', inasmuch as they may be consequences of more elementary grammatical rules, which have to be worked out logically, making a mistake, and hence discovery and knowledge possible (Schroeder, S. 2009, p. 104).

24 Wittgenstein took the coherence of our scientific world-view to be of great importance to the justification of empirical claims (OC §§102f, 126, 140-4, 603).

progress in science and mathematics, even going so far as to say that an investigation of mathematical concepts

> ...might deserve the name of an investigation of the 'foundations of mathematics'. (PI, p. 232)

The sense in which a philosophical investigation might serve as a foundation for science and mathematics, however, is not that in which traditional philosophers have generally thought it might. It isn't that we need the foundation before we can really get anything done in science, or rely on it. (MS 291, 10; Kenny 2002, p. 14) Rather, clarity in the sciences is valuable because it may forestall misunderstandings that can lead to fruitless research due, for example, to the conflation of a conceptual question with an empirical one (Bennet & Hacker 2003, p. 2). Psychology, for example, may be hampered by confusions about the criteria for the application of psychological concepts, and mathematics can go down blind allies in regard to questions about the infinite, or in pursuing confused ideas such as set theory (PI, p. 232).

The Importance of Metaethics

Similarly, then, might not an investigation of ethical concepts deserve the name of 'an investigation of the foundations of ethics'? I think that some examples of the possible impact of conceptual confusion on ethical thought indicate an affirmative answer.

Philosophical beliefs *needn't* impact upon one's life. The solipsist, for example, may treat others the same way in his ordinary life as anyone else does (Hacker 1990b, p.

503). Inasmuch as this is the case, Wittgenstein seems to have thought that metaphysical beliefs were harmless enough. Thus, in discussing a philosophical doubt about the existence of unperceived objects, Wittgenstein says

> ...how would his doubt come out in practice? And couldn't we peacefully leave him to doubt it, since it makes no difference at all? (OC §120)

Similarly, he describes the difference between a realist and an idealist who both, nonetheless, teach their children to act in the same ways (to fetch and to sit on chairs, etc.), as a mere difference in 'battle-cry' (OC §339).

But whilst a philosophical belief that makes no difference to one's life may, indeed, be innocuous, history abounds with examples of those who have not been so un-swayed by such beliefs. That Descartes (allegedly[25]) vivisected live animals (Allen & Trestman 2015, §3) with no regard for their well being in the confused belief that they were mere automata is but one particularly grizzly example.

It is easy, however, to see, more generally, how ideas based on a lack of clarity regarding the criteria for psychological ascriptions, can be put to work in, for example, providing a basis for ethical egoism, or worse. Wittgenstein once discussed the example of

> A tribe that we want to enslave. The government and the scientists give it out that the people of this tribe have no souls; so they can be used without scruple for any purpose whatever...we have also found that these people can be used successfully as experimental objects in physiological and psychological laboratories; since their reactions – including speech-reactions – are altogether those of men endowed with souls. (RPP I §96; cf. Z §528)

25 See Cottingham 1978 for a defence of Descartes on this point.

This isn't to admit that philosophy on its own can decide matters of empirical truth and falsity. Philosophy can't tell me that the people around me aren't automata, nor that they are. What philosophy tells me (or reminds me of) is that I can tell this for myself by observing their behaviour. The solipsist's confused position, though, amounts to the claim that there can be *no such thing* as other people; his isn't the empirical claim that there *aren't* any. However, the solipsist's nonsensical claim that there *could* be no other people may seem, without due reflection, to entail the empirical claim that there *aren't* any and this could lead to the confused thought that, since I am a person, but others are not, they are not worthy of the same respect as me.

Of course a metaphysical view does not *necessarily* lead to (or spring from) any particular ethical view (such a claim, would, after all, be inconsistent with the argument of the remainder of this thesis.) It can, however, be difficult to resist drawing substantive conclusions from confused metaphysical beliefs. Wittgenstein makes a similar point in relation to the dogma of the Catholic Church:

> ...dogma is expressed in the form of an assertion & is unshakable, & at the same time any practical opinion can be made to accord with it; admittedly this is easier in some cases, more difficult in others. It is not a *wall* setting limits to belief, but like a *brake* which in practice however serves the same purpose; almost as though someone attached a weight to your foot to limit your freedom of movement. (CV, pp. 32f)

But the fact that an unchecked metaphysical belief can be taken to support a moral view which one finds abhorrent provides in itself no good grounds for challenging such beliefs *tout court*. For even if abhorrent moral views can be undermined by the dissolution

of conceptual confusion, so too can benign ones. For every cold-hearted solipsist, there may, after all, be a panpsychist whose metaphysical beliefs underpin his feeling of kinship with the world and his attitude of care and respect for all things. Thus the value of philosophical reflection cannot lie in the fact that it undermines obnoxious attitudes. Alas, logic is coldly indiscriminate when it comes to the value (or otherwise) of ill-founded beliefs.

But the fact that conceptual confusion can lead us into making moral choices that we would otherwise not have made is nonetheless significant. For, as moral agents, we want to get the answers to moral questions *right*. If the remainder of what I have to say in this thesis is correct, then the attempt to get the answers to moral questions right is the attempt to live in a way that one can ultimately square with one's own conscience – one's carefully considered attitudes of approval and disapproval. But this makes getting the answers right no less important in our lives. For, to base moral judgements upon unclear thinking leaves us in danger of acting in ways that, having considered the matter more carefully, we may find difficult to live with.

These observations indicate one way in which Wittgensteinian philosophy in general can be valuable. But wherein, then, lies the specific value of metaethics? There is, of course, no general answer to the question. The importance of metaethics lies in the particular confusions with which it deals, and in their consequences. Thus some examples are called for.

As with other confused metaphysical beliefs, whilst confusions about the nature of moral judgements do not *entail* faulty substantive moral views, it is easy to see how they *might* influence one's moral outlook. It is easy to see, for example, how hard-nosed realism might underpin a less tolerant attitude towards moral differences, and equally easy to see

how stolid anti-realism might engender a 'devil may care' attitude.[26]

Wittgenstein was particularly concerned with a more subtle way in which confused philosophical beliefs might influence one's moral thinking. Wittgenstein's contributions to normative ethics were extremely sparse, and confined to the period of his work up to the completion of the *Tractatus*. One explanation for this fact is to be found in what he *did* say about metaethics. Wittgenstein took ethical standards to be largely a matter of personal inclination. His mature metaethical stance, as I shall argue, is best captured and refined in a kind of subjectivism. Thus, one point on which Wittgenstein consistently insisted in relation to normative ethics was that it 'can be no science' (see e.g. LE, p. 12; CV, p. 23; Drury 1984, p. 99). Wittgenstein was particularly concerned to dismiss the notion of an ethical 'theory'. Thus, in conversation with the Vienna Circle, Wittgenstein once said that

> If I were told anything that was a theory, I would say, No, no! That does not interest me. Even if the theory was true, it would not interest me it would not be that I was looking for. What is ethical cannot be taught. (WVC, pp. 116f)

The notion that ethics 'cannot be taught' may seem palpably false. Our basic ethical commitments are, after all, usually largely the result of a 'moral education'. We are, as children, *trained* to have certain responses to certain kinds of behaviour. Wittgenstein, however, did not wish to deny that ethics can, in this sense, be taught. What he meant by this is further elucidated (Christensen 2011, pp. 1f) in another passage, where he says:

26 Simon Blackburn once remarked that people often expect the anti-realist to have “the morals of a French gangster”! (Blackburn 1984, p. 197)

> If anyone should think he has solved the problem of life & feels like telling himself everything is quite easy now, he need only tell himself, in order to see that he is wrong, that there was a time when this “solution” had not been discovered; but it must have been possible to live then too & the solution which has now been discovered appears in relation to how things were then like an accident. And it is the same for us in logic too. If there were a “solution to the problems of logic (philosophy)” we should only have to caution ourselves that there was a time when they had not been solved (and then too it must have been possible to live and think). (CV, p. 6)

According to Wittgenstein, one can no more have an ethical theory, in the action-guiding sense that many moral philosophers have aspired to, than one can have a theory about the use of a word which would help one to decide on the correctness or otherwise of its applications. Any 'theory' about the use of a word could do no more than *describe* its use; it could not stipulate that, despite how we ordinarily use a word, it should, in fact, be used differently, nor could it adjudicate in cases where we are unsure what to say.

Similarly, an ethical 'theory' could not be 'action-guiding' in the way that some moral philosophers have wanted their theories to be. Even if, as Wittgenstein says, the theory were 'true', this could mean no more than that it correctly described one's moral standards. Such a theory can't tell us what to do in a particular case, since the standard of correctness for such a theory could only be constituted by whether or not, having considered the matter carefully, we endorse its recommendations for action.

This isn't to say that Wittgenstein was against attempting to find answers to genuine ethical problems, to tracing the consequences of adopting certain ethical principles and seeking consistency amongst our moral judgements, and even to attempting to persuade

others to adopt a particular ethical stance, by pointing out their ethical commitments (cf. Rhees 1965, pp. 21f).

The notion that there might be a moral theory, however, can lead to moral error. If one is somehow under the impression that, say, the principle of utility captures a universal truth about the world, then one may be tempted to follow the rule against opposing instincts in particular cases where, in the absence of this belief, one would have, instead, made a principled exception on the basis of the opposing instinct. That is, to take the supposed correctness of an ethical theory as a reason for action is to put the cart before the horse, for it is what we take as reasons for action that determines the theory's correctness.

It ought not to be thought, though, that the value of metaethics on this account lies in the fact that it allows moral agents to fulfil a *duty* to think clearly. Given that conceptual confusion involves a *mistake*, a moral judgement based on a conceptual confusion is also a *mistake*. As moral agents, we don't want our moral judgements to be mistaken. Hence it is imperative for each moral agent to eschew conceptual confusions. This imperative, however, can't itself be described as a moral imperative. Rather it is part of moral judgement that one attempts to judge correctly. Not to be concerned to avoid mistakes in one's moral thinking is to refuse to engage in moral thought at all.[27]

This might leave Wittgensteinian metaethics looking, once again, like a peculiarly inward-looking affair. For why, on such an account, ought one to attempt to get *others* to think clearly, especially when doing so may make them give up an attitude with which one agrees, but which for them is founded on a conceptual confusion? Ought we not to let our omnibenevolent panpsychist remain confused?

It ought, firstly, to be noted that the same objection might be brought against moral philosophy of *any* stripe. One cannot, after all, guarantee that one's arguments will be found convincing, or even that they will not lead people to take up a contrary position.

27 See Chapter Five for further discussion of this point.

Moreover, the question is not a conceptual, but a practical and moral one. Agreement founded on sand is liable to subside. We want to be able to rely upon the moral responses of others, to know where we stand with people. Hence there are pragmatic reasons why clear-headed disagreement may be preferred to muddled agreement.

More important still is the question's moral aspect, which concerns the means by which it is permissible to advance one's own moral agenda. This is a question with a long and vexed history. Suffice it to say that, for me, the notion that one should promote one's wider moral agenda at the cost of honesty and rational debate is a worrying one, and whilst much more might be said by way of support for this moral view, by the argument of this thesis, I can, in the end, only hope that the reader shares these misgivings.

The value of Wittgensteinian philosophy-as-conceptual-clarification, then, comes from many quarters. Not only is clarity an end in itself for thinking people, but philosophy can act as a salve for those who are truly troubled by philosophical problems. Moreover, conceptual clarity can help to avert real mistakes in the sciences, mathematics, and in how one lives one's life. So, whilst philosophy can be a hunt for hunting's sake, or one engaged in as a therapeutic pastime, we can also use Wittgenstein's philosophical tools to hunt for food to sustain us in our moral endeavours.

II

Metaethical Preoccupations: A Brief History of Wittgenstein's Ethical Thought

In the previous chapter, I explained that Wittgenstein made little in the way of a direct contribution to ethics. I also claimed, following remarks made by Wittgenstein himself, that we should not be interested primarily in what he had to say specifically about ethics, but in the wider application of his conception of philosophy to this area.

Whilst it is true that what is of greatest importance in Wittgenstein's philosophy is his revolutionary 'method' of philosophizing, it nonetheless stands to reason that there may be something to gain from scrutinizing Wittgenstein's own attempts to apply this method. Indeed, given that Wittgenstein's method is (strange as it may sound) the result of its own application,[28] if one takes the method to be sound, then it follows that there is at least one area in which he applied it with supreme skill.

Moreover, Wittgenstein arguably made important contributions to a number of key philosophical debates through the application of his method, perhaps most notably in the philosophy of mind, where he employed his method in the creation of the 'private language' arguments of the *Philosophical Investigations*, but also in areas as diverse as the philosophies of science, mathematics, aesthetics, religion, action and epistemology. As the progenitor of a way of doing philosophy, then, and, arguably, its most skilful practitioner in a number of areas, it would be foolish to ignore what Wittgenstein himself had to say on any topic to which he made even the slimmest of contributions.

28 Wittgenstein denied that his comments about the nature of language and philosophy constituted a 'metaphilosophy', any more than an orthographer's treatment of the word 'orthography' constituted 'metaorthography'. Whilst language does occupy a unique place in philosophy, both as subject and medium, this does not mean that its philosophical treatment requires methods different in kind to those used in relation to other areas of philosophy (PI §121).

In this chapter, then, I offer a brief history of Wittgenstein's ethical thought. Whilst, as mentioned, Wittgenstein's contributions to ethics were relatively sparse, it would nonetheless be beyond the scope of this thesis to analyse in detail every comment Wittgenstein made on ethics.[29] Instead, my aim is to identify a number of key themes in Wittgenstein's thinking on ethics, in order to guide the direction of the argument in the chapters to come.

The Notebooks and the *Tractatus* (1916-1921)

Wittgenstein's first[30] recorded comments on ethics are to be found in his wartime notebooks, dated 11th June 1916 (Lewy 2007, p. 21; Schroeder, S. 2012, p. 367). That he began to think about the place of ethics within his work on language and logic only at this late stage makes the oft-quoted claim that "the point of the book is ethical" (LF, p. 94) rather implausible,[31] if taken to mean that the *Tractatus* was predominantly a work on, or concerned with, ethics, although it is certainly true that Wittgenstein saw the ethical implications of his predominantly logico-linguistic work as its most important aspect (Schroeder, S. 2006a, pp. 29f).

The notebooks contain Wittgenstein's only contribution to first-order ethics. Here, Wittgenstein outlines a form of stoicism, heavily influenced by Schopenhauer (Schroeder, S. 2012, p. 368; Monk 1991, pp. 140-4).

Biographically, the occurrence of these thoughts at this late stage in Wittgenstein's

29 Indeed, given Wittgenstein's notoriously impenetrable style, simply identifying exactly which of his remarks were intended as contributions to ethics would be a feat of scholarship in itself.

30 Whilst some comments on Wittgenstein's personal moral outlook can be found in earlier private diary entries (GT), the Notebooks contain the first ethical writings of a distinctively philosophical nature (Lewy 2007, p. 21).

31 Indeed, this claim seems hard to reconcile with Wittgenstein's insistence that the book's 'fundamental thought' is that "there can be no representatives of the *logic* of facts" (TLP 4.0312; Wittgenstein's emphasis).

work on the *Tractatus* can be explained by the hardships he was undergoing at the Eastern Front (Schroeder, S. 2006a pp. 102ff; Monk pp. 140f). Philosophically, though, Wittgenstein's stoicism can be seen as an attempt to show how a 'good life' could be logically guaranteed.

This took essentially only two ingredients. The first of these was a conception of what 'the good life', if achievable, would be. Wittgenstein took it as axiomatic, and even "tautological" that "the happy life is good, the unhappy bad" (NB 30.7.16). He defined 'happiness' as 'being in agreement with the world' (NB 8.7.16) and appears to have had in mind that 'happiness' was something like desire satisfaction, or at least the absence of desire frustration (NB 5.7.16; TLP 6.374).

To 'be in agreement with the world' is to be in agreement with "that alien will on which I appear dependent" (NB 8.7.16). It seems that in this regard, Wittgenstein saw the happenings of the world as being, in a Schopenhauerian bent, a manifestation of a will that is independent of one's own. Thus the second ingredient of Wittgenstein's argument for stoicism was that "The world is independent of my will" (NB 5.7.16; TLP 6.373). That is, things can always go contrary to my will. Moreover,

> Even if everything that we want were to happen, this would still only be, so to speak, a grace of fate, for what would guarantee it is not any logical connexion between will and world, and we could not in turn will the supposed physical connexion. (NB 5.7.16; TLP 6.374)

Given, then, that to live a good life is for the world to be amenable to one's will, but that it is always possible for things to go other than in accordance with one's will, the only way to be logically guaranteed a good life is to make one's will agree with the world,

however it happens to be – to resign oneself to one's fate in order that one's will can never be frustrated – to cease altogether, that is, to will.

Wittgenstein however, was concerned by a number of objections to such an ethic. For one thing, he showed awareness of the fact that to cease to will is literally impossible. For, it would require one to cease to have any goals, and thus to act intentionally at all! Thus, Wittgenstein comments that someone who didn't *want* anything "would not be alive" (NB 21.7.16).

For another, he was worried that stoicism runs counter to many of our everyday moral judgements. According to common judgement, it is good, for example, to want someone else to be fortunate (NB 29.7.16). Moreover

> "To love one's neighbour" would mean to will! (NB 29.7.16)

Wittgenstein questions whether he is committed to overturning these common ethical intuitions. He responds that it is only 'in a sense' that not willing is the only good and suggests that the answer may not lie in *not* willing, but in *how* one wills (NB 29.7.16).

The idea is fleshed out somewhat in the *Tractatus*, where Wittgenstein makes a distinction between the will as a phenomenon studied by psychology and the will which is the bearer of the ethical (TLP 6.423). In the Notebooks, Wittgenstein explicitly states that "The will is an attitude of the subject towards the world" (NB 4.11.16). But the world includes the will as a phenomenon studied by psychology. Therefore it would seem that Wittgenstein's idea is that in order to live a good life is not necessary (or even possible) to renounce one's worldly, psychological will, but, rather it is necessary to alter the will that resides on a 'higher', metaphysical plane.

Indeed, in the notebooks, Wittgenstein notes that one *can't* make oneself happy

without more ado (NB 14.7.16) and that one *can't* ward off the misery of the world. The good life, he says, is that which is happy *in spite of* the misery of the world (NB 13.8.16).[32] It is not worldly misery that we are to forestall – this cannot be achieved. Indeed, Wittgenstein appears, in the Notebooks, to make suggestions which run counter to stoicism, claiming, for example, that the 'happy life' is the 'life of knowledge' (NB 13.8.16). Yet surely the 'life of knowledge' involves the striving to attain knowledge, a goal which may be, and indeed often is, thwarted, at least temporarily.

Perhaps the most charitable interpretation of Wittgenstein's stoicism, then, is that he allows that the will as a phenomenon studied by psychology may suffer disappointment, and yet this needn't lead to 'unhappiness' in the sense of having one's metaphysical will thwarted. That is, one may be disappointed, yet (somehow) not allow this disappointment to touch one's inner being. What Wittgenstein seems to have been driving at was that the good life involved taking a kind of second-order attitude towards the satisfaction or disappointment of one's worldly will. It is noteworthy in this regard that a favourite experience of Wittgenstein's in later discussions of the attitude of the religious thinker was the feeling of being 'absolutely safe' – the feeling that no matter what happened, one couldn't, in some sense, come to any harm – an expression which Wittgenstein claimed didn't make sense if taken literally, but nonetheless expressed a certain attitude, viz. that, no matter what happened to one, it was of no consequence (Schroeder, S. 2007, p. 450) (of which more later).

The notion that one can be disappointed, yet not allow such disappointments to touch one on a deeper level is perfectly coherent. It is, of course, possible, in the figurative

32 Passages of this ilk have led some commentators (Weiner 1992, pp. 101f; cf. Magee 1983, p. 287) to argue that Wittgenstein's aim was to resolve a tension often noted in Schopenhauer, between willing what is good and not willing, with the suggestion that one can will, yet not be unhappy if one's will is thwarted. Whilst the idea is hardly revolutionary, and whilst it is implausible to suppose that Wittgenstein's intention was to make a serious contribution to Schopenhauer scholarship (Schroeder, S. 2012, fn5), he does indeed appear to offer this as the solution to the problems he recognised in his own stoicism.

sense in which the poem was intended[33] to “meet with Triumph and Disaster, and treat these two impostors just the same”.[34] Insofar as such attitudes exist, however, they are no more 'other-worldly' than the disappointments which they modify.

Why, then, was Wittgenstein so intent on setting the will that is of importance to ethics 'outside the world'? In the Notebooks, we find the thought that if ethics were dependent on something *in* the world, then it would be dependent on something contingent, and thus itself contingent (NB 2.8.16).

The non-contingency of ethics is the basis for Wittgenstein's infamous claim, at the culmination of the *Tractatus*, that “ethics cannot be put into words” (TLP 6.421). Wittgenstein draws this conclusion from his principle of 'bipolarity', an implicit consequence of his 'picture theory' of meaning. According to the picture theory, a proposition serves essentially to represent, or to picture, a fact, which is a combination of the 'states of affairs' represented by the 'elementary propositions' which, according to the early Wittgenstein, are the final results of logical analysis. Thus, for the early Wittgenstein, all propositions are truth-functions of elementary propositions (an elementary proposition being a truth-function of itself (TLP 5)).

It is essential to such pictorial representations that they can represent the world either accurately or inaccurately – that the world can either be or not be as pictured by a proposition. Thus, for the early Wittgenstein, only contingent, empirical propositions, i.e. “propositions of natural science” (TLP 4.11, 6.53) are, strictly speaking, meaningful. Tautologies and contradictions are 'senseless' ('*sinnlos*'), since, whilst they are not bipolar, they are nonetheless truth-functions of elementary propositions (as shown by their truth-

33 One cannot literally treat triumph and disaster 'just the same', for one's responses on having one's desires thwarted are criteria of one's having genuinely had those desires in the first place. Were I to jump for joy (or remain entirely impassive) no matter what happened, it wouldn't be said of me that my desires had been thwarted. This is not, of course, how the early Wittgenstein would have put things, but it seems that something like a dim recognition of this fact was at play in Wittgenstein's qualms over this potential way out of the stoic's problem.

34 Rudyard Kipling – 'If' (http://www.kiplingsociety.co.uk/poems_if.htm)

tables.) Anything that does not fit into either of these categories is nonsense ('*Unsinn*') (Schroeder, S. 2006a, pp. 63ff).

Ethical pronouncements (or, at least, *some* ethical pronouncements), though, according to Wittgenstein, are not bipolar. Thus, Wittgenstein concludes, they do not represent facts in the world:

> The sense of the world must lie outside the world. In the world everything is as it is, and everything happens as it does happen: in it no value exists – and if it did exist, it would have no value. If there is any value that does have value, it must lie outside the whole sphere of what happens and is the case. For all that happens and is the case is accidental. What makes it non-accidental cannot lie within the world, since if it did it would itself be accidental. It must lie outside the world. (TLP 6.41)

Of course, not *all* ethical pronouncements are non-contingent. An action which, under one set of circumstances, might be wrong, could, in another set of circumstances, be right. The vet's putting down my hamster might be wrong if the hamster is in good health, but right if the hamster is suffering a great deal of untreatable pain. Hence the wrongness of Dr. Hill's putting down my hamster is contingent upon facts about the world (or, more specifically in this case, about my hamster.) Nonetheless, it doesn't just happen to be wrong to kill a perfectly healthy animal for no reason (cf. Schroeder, S. 2006a, p. 100). Thus the fundamental principles ('values that have value (in themselves)' (Rhees 1965, p. 17; TLP 6.422)), at least, to which we ultimately appeal in justifying particular moral judgements are not contingent upon how the world happens to be.

Less often noted is Wittgenstein's implicit explanation of the non-contingency of

ethics. Wittgenstein says that

> When an ethical law of the form, 'Thou shalt ...' is laid down, one's first thought is, 'And what if I do not do it?' It is clear, however, that ethics has nothing to do with punishment and reward in the usual sense of the terms. So our question about the consequences of an action must be unimportant. – At least those consequences should not be events. For there must be something right about the question we posed. There must indeed be some kind of ethical reward and ethical punishment, but they must reside in the action itself. (And it is also clear that the reward must be something pleasant and the punishment something unpleasant.) (TLP 6.422)

The idea seems to be that certain actions appease or offend against the will, not because of any consequences that they contingently have, but simply because they are the kinds of action that they are. That is, my attitude is an attitude towards a kind of action as such, and not towards the consequences that this kind of action might have. But that actions have their essential properties is not a contingent matter, hence it is not a contingent matter that they appease or offend against the will.

Wittgenstein appears, however, to have realised that there was a flaw in this argument. For, whilst it may not be a contingent matter that actions have their essential properties, it certainly *is* a contingent matter that one wills as one does – has the attitudes that one does towards the world. It seems, then, that Wittgenstein placed the will that is the bearer of the ethical outside the world, because if the moral status of an action is dependent upon its relation to the will, yet also non-contingent, then the will on which ethics depends must also be somehow non-contingent.

Wittgenstein's apparent concerns here are starkly reminiscent of later non-cognitivist attempts to forge a strong conceptual connection between moral judgements and attitudes, without making morality objectionably 'mind-dependent'. The problem upon which Wittgenstein had seemingly stumbled is well-known to subjectivists and their non-cognitivist opponents, namely the 'modal problem'.[35]

The obvious starting point for any ethicist who, like Wittgenstein, wants to forge a strong conceptual connection between moral judgements and attitudes is 'simple subjectivism'. For simple subjectivism claims that to say that an action is good or bad, right or wrong, is to say that one has a particular kind of attitude towards it. This neatly explains why a sincere moral judgement should necessitate an attitude on the part of the speaker. For the truth of such an attitudinal self-ascription coincides with its truthfulness (barring linguistic mistakes).[36] Thus if one sincerely says that one has a particular attitude, one does have it, just as the 'motivational internalist' claims that where one sincerely issues a moral judgement, one necessarily has the corresponding attitude.

Unfortunately, though, this also raises a problem in that it would appear to make morality objectionably 'mind-dependent'. For, if simple subjectivism is correct, then it would appear that the following is a conceptual truth:

35 This is not to say that Wittgenstein had this particular objection to subjectivism in mind here (such a claim would be implausible at best), only that there is a distinct kinship between his concerns and those raised by the modal problem.

36 This notion has, of course, come under considerable fire. One recent example of an attack on this position in the arena of ethics is to be found in Ridge 2006. It is doubtful, however, that Ridge produces any genuine counterexamples to the claim. For, in many of his examples, it is far from clear that we would say that the person concerned was mistaken, say, about his own beliefs, rather than simply displaying linguistic incompetence. Moreover, since the claim of first-person authority is not a blanket one covering *all* first-person psychological self-ascriptions, not just any counterexample will do. There are, of course, some cases in which we do not have first-person authority about our own psychology, and Ridge does, perhaps, correctly identify some of these possibilities. Thus Ridge's purported counterexamples are either incoherent, or else miss their target. (See also Chapter Three for a Wittgensteinian response to the claim that Freudian cases represent counterexamples to first-person authority).

(1) *x* is wrong iff I disapprove of *x*

We don't, however, want to say that the moral status of an action depends upon our attitude towards it. Our attitudes, we want to say, may be mistaken. We may, after all, be mistaken about the character of an action. I may disapprove of Dr. Hill's killing my hamster, and therefore say that he acted wrongly, before finding out that Squeaky had been irreparably mauled by the neighbour's cat, and that, in Dr. Hill's professional opinion as a vet, putting him out of his misery was the kindest thing to do. In such a case, we don't want to say that Dr. Hill's action was wrong until I stopped disapproving of it, but that it was right all along, and that it was my disapproval that was in error.

In order to overcome this problem, the non-cognitivist denies the subjectivist claim that '*x* is wrong' means 'I disapprove of *x*', but retains the insight that, like 'I disapprove of *x*', '*x* is wrong' functions as an expression of disapproval.

Wittgenstein's insistence on placing the will that is the bearer of the ethical outside the world, then, can be seen as an attempt to avoid any objectionably mind-dependent consequences of his internalism, much as the non-cognitivist attempts to avoid the objectionably mind-depedent consequences of simple subjectivism.

Whilst not particularly revolutionary or enlightening, Wittgenstein's ethics around the period of the *Tractatus* indicate the direction in which, as I shall argue, his ethical thought continued to travel throughout his lifetime. Two interconnected themes in particular characterise Wittgenstein's early ethics. Firstly, *'motivational internalism'* – the claim that moral judgements are internally related to conative or affective states, or attitudes – and secondly, *the non-contingency of ethics* – the claim that at least our most fundamental moral judgements are not contingent upon how the world happens to be. With this second comes *the autonomy of ethics* – the claim that moral judgements are not

themselves, nor do they follow logically from, statements of fact – since the 'truth' of a moral judgement, not being contingent upon the truth of any statement of fact cannot itself be a statement of fact. This also brings in the *mind-independence of ethics*, for, if moral judgements are not contingent upon facts, then *a fortiori*, they cannot be contingent upon psychological facts.

The Lecture on Ethics (1929-30)

Wittgenstein's only work devoted solely to the subject of ethics is his 'Lecture on Ethics', delivered some time between September 1929 and December 1930 (LE p. 3), only shortly after his return to philosophy, to a subversive anti-establishment discussion group called 'The Heretics', founded in Cambridge by C.K. Ogden.

In stark contrast to just about[37] anything else in Wittgenstein's *Nachlass*, the Lecture consists of nine pages of continuous prose. For this reason, I hope the reader will indulge a somewhat lengthy summary, since this seems the only way adequately to convey the gist of the argument.

Whilst the Lecture contains no trace of Wittgenstein's earlier stoicism, or, indeed, any first-order ethical claims, many of the metaethical points from the *Tractatus* period, as I shall argue, remain.

Wittgenstein's major new interest in the Lecture was a distinction between, on the one hand, what he called the 'trivial' or 'relative' use of value terms such as 'good', and, on the other, what he called their 'ethical' or 'absolute' use (LE, p. 5).

Of 'good' in its 'relative' sense, Wittgenstein says that it “simply means coming up

37 'Some Remarks on Logical Form' (RLF), penned at almost precisely the same time as the Lecture is perhaps the only comparable example.

to a certain predetermined standard" and that it has meaning only inasmuch as such a standard "has previously been fixed upon" (LE, p. 5). Thus, according to Wittgenstein, to call a chair 'good' means that it "serves a certain predetermined purpose", to call a pianist 'good' means that "he can play pieces of a certain degree of difficulty with a certain degree of dexterity", to say that it's 'important' not to catch a cold means that "catching a cold produces certain describable disturbances in my life" and to call a road the 'right' road means that "it's the right road relative to a certain goal" (LE, p. 5).

"Used in this way", claims Wittgenstein, "these expressions don't present any difficult or deep problems. But", he, goes on to assert, "this is not how ethics uses them" (LE, p. 5). According to Wittgenstein the essential difference between relative and absolute judgements of value is that

> Every judgment of relative value is a mere statement of facts and can therefore be put in such a form that it loses all the appearance of a judgment of value: Instead of saying "This is the right way to Granchester," I could equally well have said, "This is the right way you have to go if you want to get to Granchester in the shortest time"; "This man is a good runner" simply means that he runs a certain number of miles in a certain number of minutes, etc. Now what I wish to contend is that, although all judgments of relative value can be shown to be mere statements of facts, no statement of fact can ever be, or imply, a judgment of absolute value. (LE, pp. 5f)

Wittgenstein, then, argues that the meaning of a relative judgement of value is dependent upon the standard in accordance with which the judgement is made, but that the meaning of an absolute judgement of value is not so dependent, claiming that this feature of

absolute judgements of value is somehow problematic.

Although Wittgenstein is never clear precisely what problem this difference is supposed to bring up, it appears that he here comes close to another problem for subjectivism, namely the 'disagreement problem'.[38] When we disagree about 'relative' judgements of value, this is often a prosaic case of factual disagreement. We share the same standard of evaluation, but disagree on whether something meets that standard. We agree, for example, that a 'good car' can do nought-to-sixty in under five seconds, but have differing opinions about the performance of this particular vehicle. In other cases, we may agree on the facts, yet apparently disagree in our judgements of relative value, because we are each judging by different standards. In such a case, the disagreement can be shown to be only apparent by the recognition that different standards are at play. Once we work out that, by a 'good' car, you mean one that will do nought-to-sixty in under five seconds, and I mean one that will provide a smooth ride on a pot-holed country lane, then the disagreement evaporates.

When we come to a judgement of absolute value, however, such as a quasi-aesthetic judgement about the standards by which a car *ought* to be judged, we may not disagree with one another about any facts, but neither can the disagreement similarly be eliminated by showing our respective judgements to be based on different standards of evaluation. (cf. LE, p. 5)

Our disagreement is not one of fact, but of value, and it seems that at the time of the Lecture, Wittgenstein found this problematic. For, at this stage, Wittgenstein had no resources for dealing with anything but prosaic statements of fact, and therefore also had no resources to deal with disagreements about anything other than facts. It seems, then, that one of the “deep and difficult problems” (LE, p. 5) which Wittgenstein vaguely

38 Again, I don't, of course, mean to say that Wittgenstein had problems with subjectivism in mind in writing the Lecture, only that his concerns are at least interestingly related to such problems.

recognised as being brought up by judgements of absolute value was a problem as to what is happening when there is disagreement in such judgements – how there can be disagreement where no *facts* are contested.

At this point, then, the distinction between fact and value was still very much at the core of Wittgenstein's ethics. Thus, Wittgenstein goes on to ask his audience to imagine a book, written by an omniscient being, containing "the whole description of the world" (LE, p. 6). According to Wittgenstein

> ...this book would contain nothing that we would call an ethical judgment or anything that would logically imply such a judgment. It would of course contain all relative judgments of value and all true scientific propositions and in fact all true propositions that can be made. But all the facts described would, as it were, stand on the same level and in the same way all propositions stand on the same level. There are no propositions which, in any absolute sense, are sublime, important, or trivial. (LE, p. 6)

His eagerness to separate fact from value, however, is also still bound up with an eagerness to forestall the 'misunderstanding' that, whilst 'good' and 'bad' are not "qualities of the world outside us", they are nonetheless "attributes to (sic.) our states of mind". For, in contrast with the apparent implication of Hamlet's words "there is nothing either good or bad, but thinking makes it so" (Hamlet Act 2, Scene 2; LE, p. 6), Wittgenstein claims that

> ...a state of mind, so far as we mean by that a fact which we can describe, is in no ethical sense good or bad. (LE, p. 6)

By way of illustration, Wittgenstein invites his audience to imagine reading from the aforementioned 'world-book' a complete description of a murder, including the actions performed and the states of mind of those involved, insisting that such a description “will contain nothing which we could call an ethical proposition” (LE, p. 6).

Indeed, we might even turn our attention to our own feelings in reading such a description, or those of others who read about it, yet still, maintains Wittgenstein “there will simply be facts, facts, and facts but no Ethics” (LE, p. 7).

Wittgenstein goes on to consider what a state of affairs corresponding to a judgement of 'absolute' value would have to be, claiming that

> The right road is the road which leads to an arbitrarily predetermined end and it is quite clear to us all that there is no sense in talking about the right road apart from such a predetermined goal. Now let us see what we could possibly mean by the expression, “the absolutely right road.” I think it would be the road which everybody on seeing it would, with logical necessity, have to go, or be ashamed for not going. And similarly the absolute good, if it is a describable state of affairs, would be one which everybody, independent of his tastes and inclinations, would necessarily bring about or feel guilty for not bringing about. And I want to say that such a state of affairs is a chimera. No state of affairs has, in itself, what I would like to call the coercive power of an absolute judge. (LE, p. 7)

Wittgenstein here appears to endorse the notion that the world as we find it is 'motivationally inert'. There are no facts the recognition of which necessitate taking a particular course of action, or taking up a particular attitude. But, he claims, a judgement of

absolute value appears to say that there is such a state of affairs.

Given that the state of affairs which would have to correspond to a judgement of absolute value is a 'chimera', Wittgenstein goes on to ask

> Then what have all of us who, like myself, are still tempted to use such expressions as “absolute good,” “absolute value,” etc., what have we in mind and what do we try to express? (LE p. 7)

He says that such judgements express something similar to the feeling of “*wonder at the existence of the world*” or of “absolute safety”, and are thus used similarly to phrases such as “how extraordinary that the world should exist” or “I am safe, nothing can injure me whatever happens.” But, claims Wittgenstein “the verbal expression which we give to these experiences is nonsense!” (LE, p. 8).

If, for example, claims Wittgenstein, I say “I wonder at the existence of the world”, this is a misuse of language. For to wonder at something being the case only makes sense when we can imagine its not being the case. Thus one might wonder at a dog's being so large, because one can imagine its being smaller, but one cannot wonder at the existence of the world, since one cannot imagine it not existing. Similarly, the claim that 'I am safe no matter what happens' makes no sense, since to be safe normally means that certain things will not happen to you, yet to be 'safe no matter what happens' does not similarly serve to rule out the occurrence of any event (LE, pp. 8f).

Extraordinarily, Wittgenstein claims that

> ...a certain characteristic misuse of our language runs through all ethical and religious expressions. (LE, p. 9)

He claims that judgements of absolute value are misuses of language because they seem to be similes for relative judgements of value, yet these similes can't be given any literal interpretation. (LE, pp. 9f). Such 'similes', however, claims Wittgenstein, serve to express 'experiences' which

> ...seem to those who have experienced them, for instance to me, to have in some sense an intrinsic, absolute value. (LE, p. 10)

At this point, however, Wittgenstein hits upon a problem seemingly connected to the aforementioned 'misunderstanding' about values inhering in states of mind or experiences. For experiences are facts, but, *ex hypothesi*, facts can have no absolute value. Indeed, it is, *ex hypothesi*, nonsense to say that facts have absolute value. Thus, Wittgenstein continues

> ...I will make my point still more acute by saying “It is the paradox that an experience, a fact, should seem to have supernatural value.” (LE, p. 10)

Once again, Wittgenstein's preoccupation here appears to be with something like the modal problem – that, whilst the value of a thing is not contingent upon our response and, indeed, it is not the response that is valuable, a thing's being valuable seems nonetheless internally related to our having the response.

The similarities between Wittgenstein's preoccupations in the Lecture and those of later non-cognitivists are striking enough that it has been argued (Blackburn 1998, p. 161) that Wittgenstein here expounds a form of expressivism. Against this, it has been claimed

(Macarthur 2010, p. 89) that Wittgenstein cannot be interpreted as propounding any form of expressivism in the Lecture, because, unlike expressivists, he dismisses moral claims as nonsensical, even *essentially* nonsensical (LE, pp.11f).

It seems, however, that both positions have their merits and demerits. To claim that Wittgenstein propounded a thesis that is recognizably 'expressivist' in the current sense of the term would be obviously perverse. At the time of writing the Lecture, Wittgenstein was still wed to much of the *Tractatus* theory of language,[39] and such a theory does indeed rule out explaining the *meanings* of moral judgements in terms of their use in expressing attitudes, as does the expressivist, since, according to such a theory, meaning is a feature only of descriptive propositions.

As the juxtaposition, in the discussion, of ethical and religious language suggests, Wittgenstein's thoughts about ethical language at this stage bear a closer resemblance to his later treatment of religious belief. Contrary to popular opinion, the later Wittgenstein did not give a purely expressivist account of religious belief, but instead thought metaphysical beliefs to be essential to religion. (Schroeder, S. 2007) Indeed, as previously mentioned, in the Lecture, Wittgenstein actually characterises what a belief in absolute value amounts to (LE, p. 7). It seems likely, then, that, at the time of the Lecture, Wittgenstein still clung to the notion that the 'nonsensical' metaphysical expressions of ethics nonetheless somehow hinted at 'ineffable' metaphysical truths.

In Blackburn's defence, however, when the *Tractarian* metaphysics and theory of meaning are stripped away, what appears to be left is, indeed, a form of expressivism. For, whilst, at this time, Wittgenstein still appears to have taken ethical pronouncements to hint towards other-worldly attitudes (to “go beyond the world” (LE, p. 12)), when the insistence that significant language can only serve to describe the world is dropped, along with the

39 Indeed, 'Some Remarks on Logical Form' (RLF), in which Wittgenstein attempted to save the logic of the *Tractatus* from some well-know problems, such as the 'colour-exclusion problem' was published the same year Wittgenstein began work on the Lecture.

notion that uses of language that do not fit this mould serve to hint towards a metaphysical domain, what is left is simply that ethical pronouncements, whilst not fact-stating, are expressions, but not self-ascriptions, of attitudes.

The Cambridge Lectures (1932-3)

Wittgenstein did not give the subject of ethics any significant attention for some four years after the Lecture was composed, after which time he once again brings up the subject in a series of lectures given in Cambridge. What we have of the lectures consists of notes taken both by G.E. Moore (1955) and by Alice Ambrose (AWL), one of Wittgenstein's students.

Wittgenstein had always thought of ethics, religion and aesthetics as being closely interrelated. In the Notebooks, he had identified the problem of ethics with that of the meaning of life, and the meaning of life with God, (NB 11.6.16) and held that “Ethics and aesthetics are one” (NB 24.7.16; TLP 6.421), and in the Lecture on Ethics, ethical and religious language are given identical treatment.

Whilst the connection between these areas remains strong in Wittgenstein's mind at the time of the Cambridge lectures, claims of identity have given way to a somewhat more measured view:

> Practically everything which I say about 'beautiful' applies in a slightly different way to 'good '. (Moore 1955, p. 16)

By this point, too, Wittgenstein had read Frazer's *Golden Bough* (RFGB, p. 115) and the Cambridge lectures show that he also saw connections between ethics and ritual

practices (Moore 1955, p. 16) (of which more in Chapter Seven).

That all of these overlapping subjects were tackled together in the lectures makes interpretation challenging, since it isn't always clear whether a comment is specific to one of these subjects, or is held to be a remark on something in common between two or more of them.

Once again, Wittgenstein insists that evaluative judgements are not contingent upon psychological facts. Apart from producing counterexamples to specific analyses of value words in terms of psychological states (Moore 1955, p. 18), Wittgenstein insists that, in general, a phrase like 'The bass is too heavy; it moves too much' does not say the same as, for example, 'If it moved less, it would be more agreeable to me'. Rather, the bass' moving less is an 'end in itself', as opposed to a means to the end of psychological satisfaction (Moore 1955, p. 18).

One point which Wittgenstein was concerned to get across is that there needn't be anything in common to all things we call 'good', and that, even if we were to find something in common, this needn't be what we mean by 'good' (Moore, 1955, p. 17).[40]

Perhaps most importantly, though, other comments have suggested to some commentators that Wittgenstein, at this time, drew back from his earlier adherence to the fact/value distinction. One point that has recently received particular attention is Wittgenstein's claim, at this juncture, that words such as 'beautiful' and 'good' have different meanings in different contexts, and that this is shown by what one can say in order to convince somebody that a thing is good, or the reasons one can give for calling something 'beautiful' (Moore 1955, p. 17).

David Stern (2013, pp. 196-205) argues that this claim is strengthened by passages

40 A similar point is made by Moore in *Principia Ethica*. Moore argues that, whilst we might discover that all yellow things reflect light of a certain wavelength, it does not follow that we mean by 'yellow', 'reflects light of wavelength *x*'. Indeed, we need to know what 'yellow' means, and to be able to identify yellow things in order to be able to make the *discovery* that all yellow things share this feature (Moore 1903, p.10).

in forthcoming notes, taken by G.E. Moore, such as the following:

> Supposing you say “good is a quality of human actions & events, & one can't explain further what sort of quality”.
>
> Then ask: How does one know whether an action or event has it? (I don't despise this question: it is connected with meaning, & way (sic.) in which we learnt meaning.)
>
> Answer might be: Study the action, & you'll find out; just as you might study a thing to find out whether it's steel or not.
>
> Now can I know the action in all its details, & not know whether it's good or not? Is that it's good one particular experience, like that it's hard? Suppose I studied all the movements involved in a murder, & also all the emotions. Is there a separate investigation, having studied the whole action, whether it's good or not? (Stern 2013, pp. 203f)

Stern suggests that Wittgenstein's implicit answer is that no separate investigation is required, citing Wittgenstein's remark that the “Answer might be: Study the action, & you'll find out”, and claims that this represents a “radical departure” from Wittgenstein's adherence to the distinction between fact and value made in the *Tractatus* and the 'Lecture on Ethics'. (Stern 2013, p. 204)

Stern further backs up this claim with the following passage:

Take “elasticity”.

If I want to find how elastic a rod is, I can imagine two ways: –

(1) With a microscope I can see the structure, & can say it is elastic.

But do I mean this structure by “elastic”? I might.

But (2) I might investigate by pulling the rod, & seeing what happens.

This might be what I mean, & the structure only a symptom.

So with “good”.

We might mean by “good” simply “action of this sort”…

(Stern 2013, pp. 203f)

It appears that these passages directly target G.E. Moore's[41] claim that value words stand for a simple, unanalysable, non-natural property, and propose, instead, that value terms are simply used to describe the natural features of the object or action under consideration, and hence that the meaning of a token utterance of such a word depends upon the context, in particular on the features of the subject of the discussion which the speaker is prepared to offer as reasons for his judgement (cf. Glock 2015, p. 122).

Alice Ambrose's notes from the same lecture give a similar impression at first glance (AWL, pp. 34f). However, following on from the corresponding comments in Ambrose's notes, where Wittgenstein makes the same point regarding the relationship between 'Being beautiful' and 'Having a certain arrangement of colours and shapes', un-remarked upon by Stern, Wittgenstein adds:

41 Moore, after all, attended the lectures.

> Now no arrangement is beautiful in itself. The word “beauty” is used for a thousand different things. Beauty of face is different from that of flowers and animals. That one is playing utterly different games is evident from the difference that emerges in the discussion of each. We can only ascertain the meaning of the word “beauty” by seeing how we use it...
>
> ...in an aesthetic controversy the word “beautiful” is hardly ever used... We only use it to say, “Look, how beautiful”, that is, to call attention to something. The same thing holds for the word “good”. (AWL, pp. 35f)

These comments somewhat weaken the impression that Wittgenstein thought of such terms as having a *purely* descriptive meaning. For one thing, Wittgenstein's suggestion that “no arrangement is beautiful in itself” hints strongly that a thing's beauty is not to be identified with any of its natural properties, for he never suggested that objects didn't have *these* properties 'in themselves'. Moreover, if beauty is not a property belonging to the object 'in itself', this suggests that it involves some kind of relation, perhaps between the object's properties and our aesthetic sensibilities (cf. Z §551).

Again, Wittgenstein's insistence that the word 'beautiful' occurs only infrequently in aesthetic discussions and then serves only to “call attention to something” does not sit comfortably with the suggestion that he thought of such words as describing things in the way that a word like 'green' does. For such descriptive terms are not at all rare in aesthetic discussions and neither is their function simply to 'call attention to something'.

Even if Wittgenstein did briefly give up the distinction between fact and value as Stern suggests, though, it seems likely that this was a relatively fleeting attempt to overcome the problems that had troubled him in the Lecture on Ethics, regarding 'absolute' judgements of value. For such an approach would appear to reduce such judgements to

'relative' judgements of value – mere descriptions.

Such a solution, however, is not very plausible. Indeed, it ought to be noted that Wittgenstein's argument (if this *is* his argument) contains a *non-sequitur*. For whilst it may be true that, once a thing's natural properties are known, no further inquiry is needed in order to decide whether it is 'good', it does *not* follow that by 'good' we simply mean 'having such and such properties', any more than the fact that we decide whether a person is in pain based on his behaviour entails that by 'pain' we mean 'pain behaviour' (cf. PI, p. 179). All that the argument shows is that we decide whether or not a thing is good based on its properties, which is something that everyone will agree to.

Moreover, even if a moral judgement can serve as a description, this does not entail that the word 'good', even on a token occasion of utterance in a particular context, means the same as 'of such-and-such a description'. For it is quite possible that the word 'good' *also* serves the function of expressing an attitude in a way that such a phrase does not.

In Ambrose's notes, Wittgenstein prefaces the discussion with the remark:

> Suppose you say "*Good* is a quality of human actions and events". This is apparently an intelligible sentence. (AWL, p. 34; Wittgenstein's emphasis)

A more plausible interpretation of these passages, I submit, is that Wittgenstein is attempting to show that the intelligibility of the claim that "*Good* is a quality of human actions and events" is *only* apparent. The argument represents a dilemma for the realist. Either a further investigation is required, and the realist must explain how the claim that a thing is good is to be verified if it is not by examining its natural properties, which he is unable to do, or else no further investigation is required, in which case *the realist* has no other explanation of the meaning of 'this is good' than that it means 'this is of such and such

a description', a claim which can't withstand scrutiny of the actual use of the term 'good'.

Another prominent Wittgensteinian voice, Stanley Cavell, has argued that a conception of the use of value terms similar to that imputed to Wittgenstein by Stern overturns the 'autonomy of ethics' thesis – the claim that moral judgements do not follow logically from statements of fact.

Cavell (1999, pp. 315-9) argues that the thesis is only *prima facie* persuasive because one considers arguments such as

(2) You promised to do X

(3) You ought to do X

apart from a particular context. Apart from a particular context, concedes Cavell, the argument requires the addition of the major premise

(4) We ought to keep promises

But, claims Cavell, when we consider the argument in a particular context, the requirement to add the major premise disappears. Cavell argues for this conclusion by analogy with an argument from a different sphere. Suppose that I tell someone

(5) You ought to castle now.

and on being asked 'Why?', I reply

(6) Castling now will neutralize his bishop and develop your rook.

As the autonomy thesis has it, the inference from (6) to (5) is not valid without the addition of a major premise. But what could such a premise be? Perhaps

(7) Whenever castling will neutralize a bishop and develop a rook, then you ought to castle.

But I don't want to say this. I want to say that you ought to castle in *this* situation, not in *every* situation. But perhaps I can explain my reasoning thus:

(8) With your rook in play you can trap his queen in two moves.

(9) With his queen gone, you can win in four moves.

And (6), (8), and (9) *do* logically entail

(10) Castling now will win.

So perhaps the question is whether (5) follows from (10) (which follows from (6), (8) and (9)). But do I need to add the premises:

(11) You want to win.

and

(12) If you want to win, then you ought to do whatever will win.

in order to make (5) follow from (10)? Cavell suggests not. Indeed, he questions the intelligibility of such a premise (Cavell 1999, p. 316). In fact, Cavell claims that no major premise *could* 'fill the gap' between (10) and (5), because there is none to fill. This, claims Cavell, is because the 'content' of an 'ought' statement is *entirely* accounted for by the content of the reasons you give for it. Thus, in saying 'You ought to castle now', you have said no more and no less than that castling will win.

Firstly, it ought to be noted that it isn't at all clear that one need be able to give reasons for one's moral judgements in order for them to be intelligible. The kinds of moral judgements with which Wittgenstein was primarily concerned were, as we have seen, those directed towards a kind of action as such. But, as has recently been argued convincingly by Nigel Pleasants, such judgements are *not* based on reasons. The kinds of reasons that philosophers typically attempt to give for why killing is wrong, or why death is bad, usually at best point to analytic truths about killing and death, or at worst strike us as darkly humorous (Pleasants 2009, p. 677). They seem, at best, to say no more than that killing is wrong or death bad, because they are the kinds of things they are, which is to give no reason at all, or else simply to fail to capture the seriousness of these matters. Thus, if the content of a moral judgement is always accounted for by the reasons we can give in support of it, then it would seem that our most fundamental moral judgements are trivialities on the model of 'It's killing because it's killing'.

Even if Cavell's suggestion about the content of a moral judgement is correct, though, it isn't clear that it touches the autonomy thesis at all. For whilst Cavell may be right that, in a particular context, no major premise is required to 'fill the gap' between a factual minor premise, in the form of a reason, and a moral conclusion, or even that in some cases no such premise could do the job, the real *point* of the autonomy thesis is that there is no fact the recognition of which *on its own* logically commits one to a *particular*

moral judgement. That my values commit me to calling only certain things 'good', or to describing the things I call 'good' in a certain way in no way implies that a thing's falling under a certain description *on its own* forces me to make a particular evaluative judgement. For it is only in conjunction with the fact that I have certain values that a description of an action forces a moral conclusion. Hence the point of the autonomy thesis stands. There are no facts which necessitate a moral conclusion, only facts which, in combination with a particular evaluative outlook, necessitate a moral conclusion. But this was never denied by those who espouse the autonomy thesis.

There is, however certainly something in the notion that an evaluative utterance *can* serve as a description of the thing evaluated, although it is far from clear that it *always* does this, or that this is the *primary* role of evaluative judgements. Indeed, it would appear that the ability of an evaluation to perform such a role is in fact dependent upon its primary use as an expression of an attitude.

To illustrate the point, let us imagine a simple language-game, in which we are sorting apples by colour. All red apples are to be kept, whilst all apples of any other colour are to be discarded. The game is governed by one simple rule, viz.:

(13) All and only red apples are to be kept.

If I put a red apple in the pile of apples that are to be discarded, you might correct me by citing (13), but you might also correct me by pointing to the apple and saying:

(14) That apple is red!

In this context (14) is not just a description of an apple, but an instruction that the apple is

to be kept.

It may be suggested that the status of (14) as an instruction is not really part of the meaning of the utterance, but, according to Gricean standards, (Grice 1989, pp. 39, 44) a mere 'implicature'. For whilst, in the context of apple-sorting according to rule (13), (14) will naturally be taken as a piece of advice as to what to do with the apple, the implication that the apple is to be kept might be immediately cancelled by, for example, adding:

(15) But don't bother picking it up and putting it in the pile with the rest of the red apples. The boss isn't paying us enough to be *that* precise.

However, it isn't at all clear that Grice's suggested test for conversational implicatures is adequate. For, consider the scenario in which you utter (14), without immediately cancelling its prescriptive force by adding something like (15), and then leave the room to fetch some more apples. Whilst you're out, I get up, remove the red apple from the pile of apples to be discarded and place it with the rest of the red apples. You then re-enter the room, notice that the red apple is now missing from the pile of apples to be discarded, and exclaim 'Why did you move the red apple!?'. I should be within my rights to retort 'Because you pointed out I'd put it in the wrong pile!', and you wouldn't get away with the rejoinder that you had simply meant to remark on the colour of the apple. If this was your aim, we should say that, in this context, you had spectacularly *failed* to achieve it, and had instead inadvertently suggested that I should move the apple.[42]

The correct test, then, for distinguishing between a genuine case of utterance

42 It might, again, be said that we only assume this because to say (14) in this context is a statement of the obvious, with no apparent point, thus flouting a Gricean maxim. Wittgenstein, however, argues that, whilst one *can*, of course, make statements of the obvious, such statements only have a sense inasmuch as one *can* elicit from the speaker *some* explanation of *why* such a statement was made, if only, for example, in order to demonstrate his linguistic competence (even to himself) (OC §352). If such an elucidation were given immediately, the implicature would be cancelled, but if only done later on, I would be justified in saying that you inadvertently suggested that I had made an error with your bizarre musings.

meaning and a mere 'implicature' is not, as Grice suggested, that of 'immediate cancelability', but instead that of 'subsequent deniability'. That is, in order for the force of an utterance to be mere implicature, where a speaker fails immediately to cancel the implication of an utterance, it must nonetheless be possible, later on, for him to deny the implication without having to admit that he expressed himself incorrectly. (Schroeder, S. 2004, pp. 65-8)

Thus, in the classic case of an academic reference in which a student is described as having excellent handwriting, without any express disavowal of the insinuation that the student's work lacks academic merit, the referee could later maintain that he had simply written an uninformative reference, which passes comment only on the student's penmanship, whereas if you ask me what I thought of Margaret Thatcher, and I notice a striking looking ship sail past behind you, and remark 'She was rather lovely!', without any further explanation, whilst I might immediately cancel the implication that the former Prime Minister was admirable by saying 'That ship, I mean', if I say only later on, when reminded of my expression of esteem for Maggie, that I meant to refer to the attractive-looking vessel and not to Mrs Thatcher, I will have to admit that, in that context, I failed in my intentions, and instead paid the Iron Lady an inadvertent compliment.

Similarly, the utterance

(16) These apples are to be kept.

will, to someone conversant with rule (13), serve as a description of the colour of the apples. Thus, if, in the apple-sorting game, one's colour-blind colleague wants to know which are the red apples, one could impart this information by pointing and saying (16). One could, of course, immediately cancel the implication that these are the red apples by

saying something like

(17) The boss now says it's the green apples we're keeping.[43]

but, having failed immediately to cancel the implication in this way, one wouldn't get away later with saying that one had not said that these were the red apples.

Most games, however, are not so simple as our apple-sorting game. Within the apple sorting game, calling an apple 'red' will, logically, carry with it the implication that it is to be kept, and saying that an apple is to be kept will, logically, carry with it the implication that it is red, at least in the absence of some immediate clarification of one's statement.

Take, however, a more complex system of rules like, for example, English law. If you say, *apropos* of nothing:

(18) Tony broke the law.

– does this utterance convey anything more specific about the nature of Tony's infringement? Perhaps in the context of a courtroom, where the points of law have been decided, and we are arguing about the points of fact only, this utterance will have a fairly specific descriptive meaning. But in general, the statement will be taken to be a relatively uninformative one about Tony's relation to some as-yet unspecified act of parliament, the judiciary and, perhaps, the penal system.

Similarly, when a person's general moral outlook is well-know, his moral assertions can take on a descriptive meaning (cf. Stevenson 1944, p. 85). If I am spouting off about

43 NB – Such a response is most naturally interpreted as a response to the question 'Which apples are we keeping?', but in this context, without further explanation, this question is equivalent to the question 'Which are the red apples?'

my pinko leftie ideology, and endorse a politician as a good man on account of his policy platform you won't expect him to be a *laissez-faire* capitalist.

This isn't to say that you will garner a *complete* description of his policies from my endorsement, but my utterance will convey something about their nature, and not just by way of implicature, but as a logical implication. For I would not get away, later on, with trying to claim that I had not meant that his policies were broadly socialist. If I do so, I admit a *mistake* in the way I expressed myself in that context.

The line between implicature and logical implication isn't always hard and fast. There will be cases where we don't know whether to say I expressed myself incorrectly or not, but it seems clear that there are cases in which the descriptive content of an expression of approval or disapproval is more than merely insinuated.

None of this, however, implies that moral disagreement is not very different in character to factual disagreement. For, even where it becomes clear that we are judging by different standards, the disagreement does not disappear. Indeed, it is at this point where moral disagreement usually becomes its most ardent.

In contexts where the moral standards of the protagonists are not known to one another, a moral evaluation alone may carry no descriptive meaning. Whilst it might well be the case, in such a context, that a speaker *intends* to convey a particular description of a thing by evaluating it, or to evaluate a thing by describing it in a particular way, without the necessary background of mutually understood standards of evaluation, their intention will not translate into meaning. For, as Wittgenstein later argued, one cannot endow one's words with meaning simply by mental acts of intention. (PI p. 18, §665) Indeed, one cannot even *intend* for one's words to be understood as having a particular meaning unless one believes that convention dictates that they *can* have this meaning in this context. (Schroeder, S. 2006a, p.151) Thus only where one believes that a particular standard of

evaluation is in play can one intend to use an evaluation in order to convey a description, and even here one might yet fail in one's descriptive intentions, if the stage is not correctly set with a mutually understood background of evaluative presuppositions.

In cases where values are not mutually shared or known, though, whilst one cannot derive a description of an action from an evaluation (or vice versa), one *can* still understand, and intend, evaluations as expressions of attitudes. This, then, as R.M. Hare puts it, is still the 'primary' use of an evaluative term (Hare 1952, ch. 7).

Lectures on Aesthetics (1938)

Wittgenstein's next significant contribution to ethics came in the form of a series of lectures on aesthetics given in Cambridge in 1938 (LC, p. vii). Once more, though, Wittgenstein is clear that much of what he has to say about aesthetics applies also to ethics (LC, passim.).

Here, Wittgenstein still insists that to call a thing 'good' is not to ascribe to it a particular quality of 'goodness'. Indeed, he insists that the fact that 'good' is an adjective misleads us in this regard (LC, pp. 1, 3).[44]

Moreover, he still insists on the conceptual connection between moral and aesthetic judgements and attitudes of approval and disapproval. He notes that a good way to overcome the temptation to think of moral judgements as ascriptions of a particular quality is to look at how they are learned, claiming that

44 Casting further doubt on the naturalist interpretation discussed in the previous section.

> If you ask yourself how a child learns words like 'beautiful', 'fine', etc., you find it learns them roughly as interjections...A child generally applies a word like 'good' first to food. One thing that is immensely important in teaching is exaggerated gestures and facial expressions. The word is taught as a substitute for a facial expression or a gesture. The gestures, tones of voice, etc., in this case are expressions of approval. (LC, p. 2)

In these lectures, though, Wittgenstein seems largely uninterested in the use of 'thin' evaluative concepts such as 'beautiful' and even 'good', commenting on a number of occasions that such words are 'hardly ever used' (LC, pp. 2, 3, 5, 11).

Wittgenstein's lack of interest in such terms however, needn't indicate that he thought of such 'thin' concepts as being relatively unimportant for an understanding of ethics or aesthetics in comparison to their 'thick' cousins, as do some Wittgensteinian virtue ethicists (see e.g. McDowell 1979, 1990). Wittgenstein's nonchalance regarding such terms seems instead to stem from the fact that he saw no interesting philosophical problem with their use. To him, it seems, it is obvious that these terms are used to express attitudes of approval and disapproval in such a way that there is no sense in trying to discover who is 'right' when it comes to cases of disagreement about such terms.

For Wittgenstein, the idea that there might be a 'science of aesthetics' which would be able to settle such disagreements was

> ...almost too ridiculous for words. I suppose we ought to include also what coffee tastes well. (LC, p. 11; cf. pp. 17, 19)

Thus, when Lewy suggests that

> If my landlady says a picture is lovely and I say it is hideous, we don't contradict one another (LC, p. 11)

Wittgenstein responds

> In a sense you do contradict one another. She dusts it carefully, looks at it often, etc. You want to throw it in the fire. This is just the kind of stupid example which is given in philosophy, as if things like 'This is hideous', 'This is lovely' were the only things ever said. But it is only one thing amongst a vast realm of other things – one special case. Suppose the landlady says: “this is hideous”, and you say: “This is lovely” – all right, that's that. (LC, p. 11)

In the earlier Cambridge lectures, Wittgenstein had claimed that

> Whenever we get to the point where the question is one of taste, it is no longer aesthetics. (AWL, p. 38)

By 'aesthetics', though, Wittgenstein means not the philosophical discipline, but everyday appreciation and critical evaluation of art, literature, architecture, fashion, etc. (Schroeder, S. (Forthcoming)). One of Wittgenstein's main points seems to have been that everyday artistic controversies do not simply take the form of one person's claiming a piece is beautiful, and another ugly – simple clashes of attitudes – but are largely descriptive in nature. As Moore recalls of Wittgenstein's earlier lectures:

> The question of Aesthetics, he said, was not “Do you like this?” but “Why do you like it?”... Reasons, he said, in Aesthetics, are “of the nature of further descriptions”... He said that if, by giving “reasons” of this sort, you make another person “see what you see” but it still “doesn't appeal to him”, that is “an end” of the discussion. (Moore 1955, p. 19)

Wittgenstein's position here seems similar to A.J. Ayer's in *Language, Truth and Logic*. Ayer argued that genuine moral argument can only occur against the background of a shared set of values. Here, facts can be brought to bear in attempting to change the other person's mind (Ayer 1936, pp. 115f). Where values are not shared, we simply have opposing attitudes, and nothing more can be said (cf. OC §611; Moore 1955, p. 19).

Wittgenstein's interest at this point was instead directed towards aesthetic controversies and puzzles which stood some chance of resolution. According to Wittgenstein, 'genuine' aesthetic controversies come up between protagonists who *share* certain tastes and inclinations. There is no one absolutely correct standard of aesthetic judgement, but different genres of art and different cultures do have their own standards for artistic evaluation (although there is by no means a completely *rigid* system of rules of aesthetic evaluation even within a particular artistic culture) (Schroeder, S. (Forthcoming)).

One new facet of Wittgenstein's thought at this time was the notion of a 'cultured taste'. Wittgenstein's major point in this regard seemed to be that we do not take all evaluative judgements equally seriously, but that we take more seriously those that are informed by a 'cultured taste' – a preference characterised by a detailed knowledge of the subject matter, a (loose) adherence to certain conventional rules, and a striving for consistency in judgement. (Schroeder, S. (Forthcoming)) Whilst Wittgenstein did not explicitly connect this idea with ethics, I will discuss its possible application to this sphere

in Chapter Seven.

Conversations with Rush Rhees (1942-1945)

The most mature of Wittgenstein's recorded thoughts on ethics are to be found in records of conversations with Rush Rhees in 1942-5. Once again, the notion that ethical judgements involve attitudes is at the forefront of Wittgenstein's remarks. Rhees recalls a conversation about the ethical dilemma facing a man who must either leave his wife, in order to carry on with his research into cancer, or else stick with his wife and give up his research:

> "Such a man's attitude will vary at different times. Suppose I am his friend, and I say to him, 'Look, you've taken this girl out of her home, and now, by God, you've got to stick to her.' This would be called taking up an ethical attitude. He may reply, 'But what of suffering humanity? how can I abandon my research?' In saying this he may be making it easy for himself: he wants to carry on that work anyway. (I may have reminded him that there are others who can carry it on if he gives up.) And he may be inclined to view the effect on his wife relatively easily: 'It probably won't be fatal for her. She'll get over it, probably marry again,' and so on. On the other hand it may not be this way. It may be that he has a deep love for her. And yet he may think that if he were to give up his work he would be no husband for her. That is his life, and if he gives that up he will drag her down. Here we may say that we have all the materials of a tragedy; and we could only say: 'Well, God help you.'
>
> "Whatever he finally does, the way things then turn out may affect his

> attitude. He may say, 'Well, thank God I left her: it was better all around.' Or maybe, 'Thank God I stuck to her.' Or he may not be able to say 'thank God' at all, but just the opposite. “I want to say that this is the solution of an ethical problem. (Rhees 1965, pp. 22f)

Wittgenstein, however, qualifies his statement that this is “the solution of an ethical problem”, by saying that

> ...it is so with regard to the man who does not have an ethics. If he has, say, the Christian ethics, then he may say it is absolutely clear: he has got to stick to her come what may. And then his problem is different. It is: how to make the best of this situation, what he should do in order to be a decent husband in these greatly altered circumstances, and so forth. The question 'Should I leave her or not?' is not a problem here. (Rhees 1965, p. 23)

Wittgenstein seems to have thought of 'an ethics' as being a largely codified system of moral rules. He does not seem to have thought of the man who 'doesn't have an ethics' as bad or amoral.[45] Rather, the contrast appears to be set up in order to highlight the fact that there is no one 'correct' solution or approach to an ethical problem, but that how a person approaches an ethical problem, and what he views as the solution, will depend on his personal feelings and attitudes. One man may have a fairly explicit system of rules to guide him; another not, and this may contribute to the setting of the problem, such that one man sees a moral problem where another does not, or such that each comes to a different conclusion when faced with the same problem, but neither can be shown to be *the* 'right' or

45 Nor, as some particularists have it, does he suggest that the man who *does* have 'an ethics' is in some way falling into moral error.

'wrong' solution. Thus, Wittgenstein continues:

> “Someone might ask whether the treatment of such a question in Christian ethics is right or not. I want to say that this question does not make sense. The man who asks it might say: 'Suppose I view his problem with a different ethics – perhaps Nietzsche's' – and I say: “No, it is not clear that he must stick to her; on the contrary, ... and so forth.” Surely one of the two answers must be the right one. It must be possible to decide which of them is right and which is wrong.' “But we do not know what this decision would be like – how it would be determined, what sort of criteria would be used, and so on. Compare saying that it must be possible to decide which of two standards of accuracy is the right one. We do not even know what a person who asks this question is after.” (Rhees 1965, p. 23)

Wittgenstein was keen, however, to distance himself from any relativistic consequences of the notion that there are various ethical systems:

> “Someone may say, 'There is still the difference between truth and falsity. Any ethical judgment in whatever system may be true or false.' Remember that 'p is true' means simply 'p.' If I say: 'Although I believe that so and so is good, I may be wrong': this says no more than that what I assert may be denied.
>
> “Or suppose someone says, 'One of the ethical systems must be the right one – or nearer to the right one.' Well, suppose I say Christian ethics is the right one. Then I am making a judgment of value. It amounts to adopting Christian ethics. It is not like saying that one of these physical theories must be

the right one. The way in which some reality corresponds or conflicts-with a physical theory has no counterpart here.

"If you say there are various systems of ethics you are not saying they are all equally right. That means nothing. Just as it would have no meaning to say that each was right from his own stand-point. That could only mean that each judges as he does." (Rhees 1965, p. 24)

Wittgenstein also insists at this point that moral judgements express rules (Rhees, 1965, pp. 23f), a suggestion which appears to have influenced some Wittgensteinian ethicists to claim that moral judgements express *grammatical* rules (see e.g. Arrington 1989; Phillips & Mounce 1969).

These 'conceptual relativists' argue that the standards of correctness for the application of a moral concept are public, shared and known within a particular moral culture, such that there are objective standards of correctness for moral judgements made within that culture, as set down by the rules for the use of the particular moral concepts at play there. Different cultures, with different moral standards, however, employ different moral concepts, such that when there is cross-cultural moral disagreement, the protagonists are talking at cross-purposes.

The conceptual relativist, then, appears to liken those with different moral concepts to those with different colour concepts.[46] There could, after all, be beings with different perceptual capacities, who are able to see colours which we cannot, and hence to use different colour concepts.

This approach is *prima facie* attractive, since it seems to capture many of the features of moral discourse which Wittgenstein thought problematic, maintaining a

46 In this respect, 'conceptual relativism' is not dissimilar to the anti-anti-realism of McDowell, Lovibind e*t al*.

conceptual connection between moral judgements and attitudes, distinguishing between facts and values, making values ultimately non-contingent and mind-independent, and apparently explaining the intractability of moral disagreement.

Wittgenstein, however, made no such claim, and indeed, 'conceptual relativism' appears to have consequences which Wittgenstein expressly disavowed. In particular, a problem arises out of the aforementioned analogy, on such an account, between cases of divergent moral concepts and cases of divergent colour concepts. For divergence in colour concepts straightforwardly implies the kind of relativism that Wittgenstein was keen to deny in the case of ethics. If there were 'bee people' who could visually perceive light in the ultra-violet spectrum, and described, say, buttercups, using a colour concept unusable by us – calling buttercups 'ultraviolet', for example – then their descriptions would usually be quite straightforwardly *true*, just as our descriptions of these flowers as 'yellow' are usually true. Neither would it be the case that they are 'yellow-for-us', but 'ultraviolet-for-them', except inasmuch as this means that we perceive one colour where they perceive another (cf. Hacker 1987, p. 152). Once our divergent standards of assessment are unearthed, the disagreement ought to disappear in the moral case, just as it would in the case of colours. This, however, does not happen.

The conceptual relativist might respond that, what makes the concepts of each culture distinctively moral is their role in expressing attitudes of approval and disapproval, hence there remain disagreements in attitude even between people of different cultures, even though there is no substantive disagreement (cf. Arrington 1989, p. 264).

Unfortunately, this response still does not overcome the fact that there is still a striking disanalogy here with the case of colours, where, when we recognise that different conceptual schemes are at play, we will happily admit that judgements which apparently contradict our own are true, in fact as true as ours (even if we are unable, due to our

differing perceptual capacities, fully to understand these judgements and the concepts they involve (cf. RC §§13, 77, 112, 120, 128)).[47]

Moreover, this account doesn't seem capable of capturing the ubiquity of moral disagreement. For even within a fairly cohesive moral culture – a particular religion, say – there can be disagreement on fundamental points. Catholics, for example, may disagree amongst themselves as to the moral status of capital punishment. In order to account for this fact, the conceptual relativist resorts to the notion that our moral concepts are 'vague'. That is, there are cases of disagreement in which the conceptual relativist does not want to say that the parties are so divergent in their moral outlooks that they are using different moral concepts, but that a single moral concept allows for a certain amount of indeterminacy, such that people sharing the concept may nonetheless disagree as to its application (Arrington 1989, p. 252).

Moral disagreement, though, rarely appears to conform to this model. For, whilst where a concept is vague, there may well be disagreement about where a line which is not drawn for us by the concept ought to be drawn for a particular purpose, there will (if the concept really is vague) be agreement that such a sharpening of the concept is an artificial stipulation, and that there is no 'correct' place to draw the line, which is not the case with most (if any) moral disagreements.

Unfortunately, Wittgenstein is never clear about the sense in which moral judgements express rules. His remarks about rules in general, however, indicate that in conceiving of moral judgements as rules, Wittgenstein remained faithful to much that he had said previously about ethics.

One aspect of rules which Wittgenstein repeatedly insisted upon was their 'arbitrariness' (PG, p. 164; Z §331; PI §497). Rules are not responsible to reality for their

47 Indeed, Arrington (1989, p. 264) claims that people using different concepts will be *unable* to call the moral judgements of those utilising different moral concepts 'true' or 'false'!

'correctness'; they are stipulations, rather than discoveries. To take a moral proposition as a rule, then, involves not allowing it to be either falsified or confirmed by experience – taking it not to be a contingent matter. Hence it appears that Wittgenstein always maintained the claim that ethical judgements, at least those of fundamental principle, were non-contingent.

This also hangs together with a distinction between rules and statements of fact, or descriptions. A rule, such as a grammatical rule, might function as a standard by which representations of facts are judged correct or incorrect, but it is not itself such a representation (Baker & Hacker 1985, pp. 262-70). Such rules are cited in order to correct linguistic mistakes and point out absurdities and patent falsehoods. They are not treated as something to be tested, but as that against which other things are tested.[48] Thus, it would appear, Wittgenstein also cleaved to the distinction between fact and value.

Wittgenstein's Non-cognitivism

Several themes, then, persisted throughout Wittgenstein's musings on ethics. Firstly, *'internalism'* or the conceptual connection between moral judgements and states of will or attitudes. Secondly, the *autonomy of ethics* – the claim that moral judgements neither are, nor do they follow from judgements of fact. Thirdly, the *non-contingency of ethics* – the claim that at least some moral judgements are not treated as susceptible of confirmation or disconfirmation by experience. And lastly, the *mind-independence* of ethics – the claim that the 'truth' of a moral judgement is not contingent upon my attitudes towards it, or the ethical system which I happen to adopt.

48 The status of a proposition as a rule, however, depends upon the context, and what is in one context treated as a rule may, in another, be subject to testing, and *vice versa* (OC §98).

Of the various metaethical options on the table, one stands out as particularly apt to capture Wittgenstein's position on ethics, namely non-cognitivism (cf. Glock 2015; Blackburn 1990, 2008; Schroeder 2006a, p. 13). Non-cognitivism was conceived with the specific intent of accounting for internalism, and to overcome the difficulties with simple subjectivism that also appear to have been close to Wittgenstein's concerns about his own internalism, not least the problem that it appears to make moral judgements objectionably 'mind-dependent'. Moreover, it achieves this, like Wittgenstein, by invoking a distinction between fact and value.

Whilst a form of non-cognitivism or expressivism may be the most natural extension of Wittgenstein's account of ethics, though, I shall argue that it would be a mistake simply to follow Wittgenstein's lead in this regard. For, so I shall argue, using Wittgenstein's broader conception of philosophy and in particular his account of meaning, we can take advantage of the virtues of non-cognitivism, without falling prey to its vices.

III

Psychological Observations – Nuancing Wittgenstein's 'Expressivism'

Before turning to this task, though, it is worth looking more closely at an area in which Wittgenstein said more along non-cognitivist lines. For my intention is not simply to disregard Wittgenstein's non-cognitivism as failed or useless, but instead, perhaps, to expand upon and to nuance it somewhat in light of his wider philosophical thought.

Although the single most important aspect of Wittgenstein's work is undoubtedly his thoughts about the nature and role of philosophy, he is perhaps better known for his contribution to the philosophy of mind. His notoriety in this regard, however, is often tantamount to infamy. Indeed, his philosophy of mind is often dismissed as a form of behaviourism, a view usually thought to be absurdly flawed due to its apparent elimination of psychological states in preference for mere behaviour, or behavioural dispositions.

It is not surprising that Wittgenstein's position has not gained widespread assent, for it runs in stark contrast to assumptions underlying positions of nearly every stripe since at least Descartes. Moreover, Wittgenstein's position is not as easy to pin down as most of his opponents would like. For, as well as being extremely nuanced, it is also not univocal in its treatment of psychological utterances, requiring, for example, different treatments of first-person psychological utterances and their second and third-person counterparts.

Whilst it is the 'behaviourist' account of utterances in the second and third-person which is most often discussed (and dismissed), of more interest for our present purposes is Wittgenstein's 'expressivist' account of certain first-person psychological utterances.

Metaethical positions which take moral judgements to be expressions of attitudes are often pilloried as 'Boo/Hurrah!' theories. Whilst technical qualms about the 'embedability' of such expressions remain at the forefront of objections to such theories,

another, related misgiving often appears to stand close behind. For the account of expressions of attitudes as akin to such simple interjections as 'Boo!' and 'Hurrah!' seems to ignore the huge variety of ways in which moral or evaluative language is used (cf. LC, pp. 2f) and gives the impression that the expressivist must think of moral and other evaluative judgements as rather primitive responses – a kind of letting off steam when the pressure gets too much (Hare 1952, p. 10). A cry of pain, or a grunt of disgust might fit this model well, but it cannot seriously be maintained that our careful, thoughtful, use of moral language is always, or even often like this.

Moreover, the expressivist's insistence that certain psychological self-ascriptions are not descriptions of one's psychological states seems counterintuitive at best. For we very often *do* describe our sensations, feelings and attitudes.

Wittgenstein was alive to such misgivings, and, indeed, took great pains to allay them. His writings, on the subject, however, have not been much discussed, and seem sometimes to have been somewhat misunderstood. In this chapter, then, I aim to outline Wittgenstein's 'nuanced' expressivism about psychological utterances, and to discover what may be gleaned from it about moral language.

The Traditional Conception of the Mental

What I call the 'traditional conception of the mental' is an assumption underlying a great deal of the philosophy of mind, going back to at least Descartes. According to this conception, mental phenomena, such as sensations, emotions, attitudes and beliefs are 'inner objects' of which one can attain knowledge through 'introspection'. On this model, psychological self-ascriptions constitute descriptions akin to those of scenes before one's

eyes. One 'looks' into one's 'inner world' and can then describe it as containing a pain, a desire or an intention, just as one can look at one's room and describe it as containing a table and chairs (PI §290).

Wittgenstein famously brought powerful objections to bear against this notion in the 'private language arguments' of the *Philosophical Investigations*, and elsewhere. The idea that psychological self-ascriptions might be based on observation of the private objects of an inner world entails the incoherent notion that there could be a 'private language' whose terms refer to what could only be known to a given speaker. One cannot give oneself a private ostensive definition of a word which is to refer to some object with which others cannot possibly be acquainted, for there could be no standard of correctness for the application of such a word. Whatever seems right to me regarding its application will be right, thus there could be no ground for saying that what I called 'pain' on one occasion was the same as what I called 'pain' on another occasion, and this, in turn, would make it impossible to teach someone else to use the word as I do (PI §258). But we understand each other perfectly well when we talk about 'being in pain', for example. Thus even if the existence of some alleged 'inner object' is assumed, the nature of this object is entirely irrelevant to the language-game with sensation-words (PI §293).

Words for mental phenomena are not used to describe so many objects of an infallible faculty of 'inner perception'. Whilst we might indeed say that we are 'aware of' or that we 'experience' our psychological states, any analogy with our awareness or experience of physical objects is misleading. In gaining knowledge of the world through perception it makes sense to improve the conditions of perception, to take a look from a different angle or in different light, to check one's observations using different senses or by using special instruments where observations cannot be made directly, to ask other people who are better suited for making a certain observation and to consult experts. None of

these makes sense in the case of one's pain.

It is, in fact, nonsense to say 'I know that I am in pain', if 'know' is interpreted epistemically.[49] Whilst the Cartesian conception of the mental is correct in noting that doubt and error are inconceivable with regard to one's own sensations, beliefs and the like, the very incorrigibility of our purported knowledge of the objects of the mental sphere shows this not to be a case of knowledge at all (PI §246; LW II, p.92). It makes sense to claim to know only when it would also make sense for one *not* to know, to find out, to check and to offer *grounds* for one's knowledge-claims, but none of these makes sense in the case of one's own current sensations, emotions and such like (PI, p. 221; cf. OC *passim.*). For one does not ascribe 'mental states' to oneself on the basis of any grounds or *criteria*,[50] as one ascribes them to others.

One cannot justify a claim to know, for example, that one is in pain by citing the fact that one *feels* pain, as one might offer as a ground for the claim that there is a chair in the room the fact that one can *see* it, for to feel pain *is* to be in pain, thus one's justification comes to 'I know I am in pain because I am in pain, i.e. because my original utterance "I am in pain" was true, that is, I was not lying' (LSD, p. 13) which is to offer no ground at all, but simply to reassert that one is in pain (Hacker 1990a, pp. 85ff).

49 One can, of course, say this in order to highlight the grammatical point that it is senseless to say that one doubts whether one is in pain (PI §247), or in order to assure someone that one is using the word 'pain' correctly (cf. PI §381; PI, p. 221).

50 A criterion is not a necessary or sufficient condition for the application of a concept. Rather, where *x* is a criterion of *y*, it is part of the meaning of *y* that *x* is good, but defeasible evidence for *y*. We ascribe 'pain' to another person by observing the evidence of his being in pain in his behaviour. But it has not been *discovered* that certain forms of behaviour accompany pain; we can only judge what accompanies pain (brain states, for example) because the concept 'pain' allows us to ascribe pains on the basis of behaviour to begin with (Schroeder 2006a, p.209 fn31).

Wittgenstein's arguments, however, are not entirely negative. In place of the traditional conception of the mental and of psychological language, which conceives of the language-game with psychological self-ascriptions as beginning with observation of a logically private 'inner object', Wittgenstein suggests instead that this language-game begins with natural expressive behaviour in certain circumstances (PI §290; Hacker 1990a, p. 88). When a child injures himself and cries, we do not doubt that he is in pain and it would be absurd, for example, to suggest that the smile of an unweaned infant was a mere pretence of contentment. Pretending involves motives and intentions the ascription of which to an infant would be absurd, since, once again, their behavioural repertoire does not include the *criteria* for such motives or intentions (PI §§249f; PI pp. 228f).[51]

The function of first-person psychological utterances, Wittgenstein suggests, is not to describe 'private objects', but to replace and extend natural expressive behaviour:

51 Wittgenstein concedes (PI §250) the possibility that an animal could be taught to act as though it were in pain when in fact it was not, but denies that this would be real 'pretence'.

Indeed, animals *do* sometimes learn, even without being taught, that they will get attention if they behave in certain ways, as when a dog, having gained attention in the past for limping, begins to limp in order to get attention. I am not so confident that we could not legitimately call this 'pretence', given that the dog has motives and intentions, which it displays in its behaviour. If a dog ceases to limp when it becomes clear that he is not going to get a treat, for example, we might say that he was pretending to be in pain.

Of course, it is arguable that animals do not have the *concepts* of pretending, or of that of which their behaviour is a pretence, which, it seems, is what Wittgenstein thought requisite for the ability to pretend. (It seems Wittgenstein thought that one who did not understand what 'pain' meant could at most *believe* that he was pretending to be in pain (PI, p. 229)). However, some adult human beings lack such concepts too and one would, I think, hesitate to say that they could not, on that account, pretend.

An unweaned infant, on the other hand, may, for example, cease to cry out when it has been punished for doing so, although it is certain that they have not thereby ceased to be hungry or uncomfortable. In this case, however, it seems absurd to suggest that the infant is 'pretending' not to be hungry or uncomfortable, as they are not yet capable of having the requisite kinds of motives and intentions. Their behaviour is strictly instinctual – an evolutionary survival mechanism. That is, it is explained *causally*, and not in terms of *reasons*. (For a compelling argument for interpreting animals as acting for reasons see Glock 2000).

> How do words *refer* to sensations? – There doesn't seem to be any problem here; don't we talk about sensations every day, and give them names? But how is the connexion between the name and the thing named set up? This question is the same as: how does a human being learn the meaning of the names of sensations? – of the word "pain" for example. Here is one possibility: words are connected with the primitive, the natural, expressions of the sensation and used in their place. A child has hurt himself and he cries; and then adults talk to him and teach him exclamations and, later, sentences. They teach the child new pain-behaviour.
>
> "So you are saying that the word 'pain' really means crying?"— On the contrary: the verbal expression of pain replaces crying and does not describe it. (PI §244)

When one injures oneself and cries out, it would be absurd to say that one is observing one's pain and then describing it. A cry of pain is 'wrung from us' when we are in pain (LW II, p.14; PI §546); it is an involuntary manifestation of one's pain, not issued with the intention of describing anything, or with any other intention for that matter (Z §48; LW I §§20, 37).[52] I simply hurt myself and cry out in pain and it will not be said in such circumstances that I have described myself or my pain.

Psychological Descriptions

Put thus, however, Wittgenstein's account may seem a fairly crude form of expressivism.

52 If I intend simply to cry out (and not, say, thereby to get you to stop what you're doing which is hurting) then I am 'putting it on'.

For the assimilation of all uses of utterances such as 'I am afraid' and 'I am in pain' to shrieks and howls is clearly implausible – as implausible as the assimilation of all moral judgements to shouts of 'Boo!' and 'Hurrah!'.

In fact though, whilst, in the sections of the *Investigations* constituting the 'private language arguments', Wittgenstein pays little heed to such concerns, elsewhere, most notably in section IX of Part II of the *Investigations* (PI II, ix), amongst other places (RPP I & II; LW I & II, Z), he gives the matter considerable attention.

David Macarthur has recently cited these remarks in outlining a reading of Wittgenstein according to which he cannot usefully be thought of as an expressivist of any traditional kind. Macarthur argues that, whilst the expressivist correctly interprets Wittgenstein as rejecting the descriptivist's notion that all indicative language serves the purpose of description, the expressivist is, like the descriptivist, still in the grip of a similar tendency to smooth over differences in language. According to Macarthur, the descriptivist is committed to four theses:

> 1) The ***functional univocity thesis***: the claim that the target discourse functions in just one distinctive way.
> 2) The ***functional transparency thesis***: the claim that a sentence's function can be simply read off from the form and/or content of the sentence itself.
> 3) The ***descriptive function thesis***: the claim that the function of the target discourse is to describe its own distinctive region of reality.
> 4) The ***unity of description thesis***: the claim that description is a single functional category (say, truth-assessable language). (Macarthur 2010, p. 86)

Macarthur argues that the expressivist rejects only the 'descriptive function thesis', whilst retaining the other three theses. In metaethics (as opposed to the philosophy of mind), the expressivist retains the 'functional univocity thesis' because he holds that the function of moral language is *always* to express an attitude, rather than to describe something. The 'functional transparency thesis' is retained, because, whilst the expressivist rejects the descriptivist notion that the descriptive function of an utterance can be 'read off' from its indicative form, he nonetheless takes it that the expressive function of an utterance can be read off from the appearance of certain key terms, such as 'good'. The 'unity of description thesis' is retained, because the expressivist cleaves to the descriptivist idea that description is always a matter of representing the facts either accurately or inaccurately.

Macarthur, however, insists that a close reading of Wittgenstein, in particular PI II ix, reveals that he rejects all four of the above theses and that he therefore falls neatly into neither the descriptivist nor the expressivist camp (Macarthur 2010, p. 86).

According to Macarthur, in claiming that psychological self-ascriptions are uniformly *not* descriptions, the expressivist errs. Instead, Macarthur appears to suggest, we should view 'description' as a something of a 'mixed bag' concept, which can be used in various different applications.

Whilst I think that there is much truth in this, it seems to me that Macarthur's emphasis on faithfulness to the everyday motley of uses to which the term 'description' is put misses an important aspect of Wittgenstein's point. For it seems to me that the point of Wittgenstein's discussion is *not* to say that we should remain faithful to the everyday concept of 'description', but that the motley nature of our ordinary concept is misleading in the context of a philosophical discussion, to the extent that, in such a discussion, we may well be better off marking different uses of the term with a difference in terminology.

Before getting on to Wittgenstein's discussion of 'description', though, it is worth

examining some of Macarthur's arguments for his liberal view of 'description'.

Oddly, one of Macarthur's main objections to interpreting Wittgenstein as rejecting the expressivist claim that psychological avowals do not serve as descriptions of psychological states is that

> ...if an avowal were a mere expression it seems that then there would be a genuine state of affairs in the world, namely, one's current state of mind, that others can describe or talk about but that I, mysteriously, could not describe or talk about (unless I'm reflecting upon it as in the past). To suppose that Wittgenstein is committed to *that* bizarre view seems most uncharitable. (Macarthur 2010, p. 84)

As we shall see in our discussion of Wittgenstein's treatment of Moore's Paradox in the next chapter, however, such an objection is rather ironic. For this was *precisely* Wittgenstein's position:

> The difficulty becomes insurmountable if you think the sentence "I believe..." states something about the state of my mind. If it were so, then Moore's Paradox would have to be reproducible if, instead of saying something about the state of one's own mind, one were making some statement about the state of one's own brain. (RPP I §501)

Moreover, Macarthur takes Freudian cases, in which, for example, one is convinced by a therapist that one is subconsciously afraid of one's father, to constitute counterexamples to the notion that psychological self-ascriptions are not descriptions of

psychological states, in that one's concession, on the therapist's couch, that one *must*, after all, be afraid of him is *not* an avowal of fear, inasmuch as one's fear is still repressed (Macarthur 2010, p. 85).

Again, given Wittgenstein's stated views on Freud, this is a surprising tack to take. Macarthur is surely correct that we would not call such an admission an avowal of fear, but for Wittgenstein this is because 'unconscious' or 'repressed' fear is not simply a species of what we ordinarily call 'fear', but something which, perhaps, is closely analogous to normal cases of fear (cf. PG, p. 48; BB, pp. 22f, 55-8; PI §149).

Macarthur suggests that Wittgenstein's refusal to accept the functional univocity thesis is in part due to the truth-evaluability of psychological self-ascriptions and their embedability within more logically complex constructions, in contrast with such interjections as 'Ouch!' and 'Hurrah!'.

Whilst it is certainly true that this distinguishes the use of such self-ascriptive sentences from the use of interjections, it is far from clear that this entails, as Macarthur claims, that these embedable constructions have a 'descriptive aspect'. Indeed, Macarthur accuses the expressivist of equating having a descriptive function with being truth-evaluable (Macarthur 2010, p. 85), but then goes on apparently to do just that!

It seems, then, that we are left with little in the way of motivation for accepting Macarthur's liberal view of description.

Philosophical Investigations Part II Section ix

Wittgenstein's actual position with regard to the use of the word 'description' is, I shall now try to demonstrate, somewhat ambivalent. Whilst, as Wittgenstein acknowledges,

Macarthur's view may well be true of the term in our common parlance, the moral of Wittgenstein's discussion is that this is a highly *misleading* feature of our language. Thus, in the context of a philosophical discussion, at least, we are better off distinguishing carefully between different uses of the term, in order to avoid conceptual confusion, even to the extent that it may be better not to use the term 'description' at all for some things of which we would ordinarily use it without hesitation.

As already noted, Macarthur's reading of Wittgenstein is based mainly on PI II ix. A close reading of these comments, however, taking into account other texts in which Wittgenstein discusses the same topic, points towards an interpretation rather different from Macarthur's.

The opening comment of PI II ix, constitutes a *coda* to Wittgenstein's earlier observations that psychological avowals are not based on observation of an 'inner object':

> If you observe your own grief, which senses do you use to observe it? A particular sense; one that *feels* grief? Then do you feel it *differently* when you are observing it? And what is the grief that you are observing—is it one which is there only while it is being observed?
>
> 'Observing' does not produce what is observed. (That is a conceptual statement.)
>
> Again: I do not 'observe' what only comes into being through observation. The object of observation is something *else.* (PI, p.187a)

If one has to 'observe' one's grief in order to feel it and therefore to know that one has it, then it would seem to follow that one could have grief that was not observed and

therefore not felt. But one cannot grieve without feeling grief, for to feel grief *is* to grieve.[53] Furthermore, the grief that one *has* and observes would have to be distinct from the feeling of grief which arose from one's observation, for if one only has (feels) grief in virtue of observing one's grief, then clearly no observation has occurred after all, for it is a grammatical truism that observation does not produce what is observed.[54] If the grief that one has is different from the grief that one feels in virtue of observing it, however, then it would still appear that one would have failed to observe one's grief, just as one has failed to observe some phenomenon in the physical world if one's observation changes what is observed (cf. RPP I §690).

In another strand of argument in PI II ix Wittgenstein applies an aspect of his earlier arguments against 'referentialism': the idea that words are names of objects and that the meaning of a word is the object to which it refers (PI §1):

> We ask “What does 'I am frightened' really mean, what am I referring to when I say it?” And of course we find no answer, or one that is inadequate.
>
> The question is: “In what sort of context does it occur?”

> I can find no answer if I try to settle the question “What am I referring to?” “What am I thinking when I say it?” by repeating the expression of fear and at the same time attending to myself, as it were observing my soul out of the corner of my eye. In a concrete case I can indeed ask “Why did I say that, what did I mean by it?”— and I might

53 This is not to say that grief is a 'feeling' in the sense of 'sensation'. One cannot feel grief for just a second, as one can pain, for example (PI, p. 174b-c).

54 One might object that observation *can* produce what is observed, just as when an experimenter, by his presence, influences the objects of observation. However, the experimenter must be able to *observe* the anomalous effects caused by his observation, thus it is not, in the sense required, his observation of these effects which causes them.

> answer the question too; but not on the ground of observing what accompanied the speaking. And my answer would supplement, paraphrase, the earlier utterance. (PI, p. 188))

It is, of course, quite natural to say that 'pain' is the name of a sensation, that 'pain' refers to pain (and not, say, to pain behaviour), or that, by 'pain' one meant (was referring to) pain. Although usually harmless, such phraseology can be misleading. What we call 'the name of a sensation' is markedly different from what we call 'the name of an object' (LSD, p.11), for, as Wittgenstein comments elsewhere:

> “The word 'pain' is the name of a sensation” is equivalent to “'I've got a pain' is an expression of sensation”. (Z §313)

Similarly, 'to refer to a sensation' is not at all the same as 'to refer to an object'. In the case of an object, one can point to the object, hang a label on it or show it to other people. None of these, however, makes sense in the case of 'pain' (Hacker 1990a, p. 91).

Although one cannot answer the question what one meant by the words 'I am afraid' or 'I am in pain' by, for example, pointing to the thing one was referring to, this does not mean that one cannot say what one meant, for the question was not a request for an ostensive definition, but for clarification, which is generally given in the form of a paraphrase of one's original utterance. This word has a use in our language that is not dependent on the possibility of pointing to a sample of pain, but is instead dependent upon the possibility of pointing to certain kinds of behaviour in certain circumstances, or eliciting from someone an expression of pain (by, for example, pricking him with a pin) (PI §288).

A further aspect of Wittgenstein's attempts to cast doubt on the notion that present tense psychological self-ascriptions uniformly constitute 'descriptions of one's state of mind' is touched upon only very briefly in PI II ix:

> What is fear? What does “being afraid” mean? If I wanted to define it at a *single* shewing—I should *play-act* fear.
>
> Could I also represent hope in this way? Hardly. And what about belief? (PI, p. 188)

Here, Wittgenstein alludes to the fact that a 'state of mind' has 'genuine duration': it begins and ends at a particular time, can be continuous or intermittent (Z §§71-85; cf. Baker & Hacker 1980a, p. 13). Wittgenstein here proposes a test which can be applied in order to distinguish genuine mental states from other mental phenomena. It is possible, for example, to imagine someone who always manifested his fear. Perhaps he would quake, or shrink into a corner whilst afraid and only cease to do so once his fear passed. We could 'play-act' being such a person. Of course, one *can*, 'play-act' believing something, but not in the same way. There is no characteristic behaviour constituting a criterion of belief; rather, one manifests countless beliefs in all of one's intentional actions. One could, of course, 'play-act' belief simply by making an assertion, but 'belief', unlike 'fear' has duration independently of its possible manifestation (PI, pp. 191f). We could not imagine a person manifesting his belief when and only when he had it, as he might his fear. Our beliefs do not come and go with consciousness, for example, as does our fear.

Wittgenstein's point here is that many psychological self-ascriptions are not even *candidates* for being descriptions of 'states of mind', since no 'state' is denoted by many

psychological terms; that is, they do not serve as manifestations of any state.

The temptation to conceive of psychological self-ascriptions as descriptions is, perhaps, at its strongest in the case of genuine states, such as pain or fear.[55] Because these states have genuine duration, we think of them as being 'there' in our inner world at a particular time to be observed and described.

In the case of many psychological phenomena, however, this analogy with objects of perception is far less plausible. Belief, intention and hope, for example, are markedly different from sensations or feelings (cf. RPP II §45).

If I say 'I'm having a party next Saturday, I hope you'll come', this second clause is not a description of my state of mind. For one thing, it is not a 'description' at all, but simply an expression of a wish. I might equally have said 'I'm having a party next Saturday, do come' and nobody would call *this* a description (RPP I §469). Neither, however, is it an 'expression' of a state of mind, for 'to intend' is *not* to be in a state of mind. If I say 'I'm having a party on Saturday, I hope you'll come', the question 'How long have you been hoping?' would usually be out of place and would in any case be irrelevant to the question of whether I do hope that you'll come. I needn't have given the matter a second's thought beforehand, nor need it occupy my thoughts thereafter (LW I §2; Z §78). The utterance, like an invitation, is simply an expression of a wish for you to come. To feel hopeful, on the other hand, can be said to be a state of mind (cf. RPP I §461). If Saturday comes around and I am waiting expectantly by the door, I may explain myself by saying 'I'm hoping that he will arrive soon' (RPP I §465).

Similarly, 'intention' is not a state of mind (Z §45) and 'I intend' is, therefore, never

55 Of course *some* cases of fear are more like belief. I may be afraid that I will lose my job or be unable to keep up the mortgage repayments and this kind of fear is clearly different from that where I am afraid of a man with a gun in front of me. The former kind of fear is more a kind of disposition, which informs my actions over a long period of time.

a description of a state of mind (RPP I §599; RPP II §178).[56] Whilst my 'thinking and acting with such and such an intention' (RPP I §598) may have genuine duration, the intention itself does not come and go along with my thinking and acting in accordance with it.

'Belief', is more akin to a disposition (PI, pp.191f; cf. RPP II §45)[57], which persists even through unconsciousness and involves no characteristic feelings or sensations. Of course, one may say that one feels strongly that such and such is the case, but this is simply to say that one is confident, but cannot produce any specific grounds for one's confidence (Schroeder, S. 2006b, p. 167). But 'confidence in one's belief' is not a feeling; to have confidence in one's beliefs is to be willing to act on them and unwilling to give them up without sufficient evidence.

It may be objected that self-ascriptions of belief are therefore descriptions of one's dispositions. However, Wittgenstein's discussion in this regard overlaps with his discussion of 'Moore's Paradox', much of which is devoted to exposing the problems associated with conceiving of a self-ascription of belief as a description of one's disposition, rather than as a manifestation of it (see PI II, x; cf. Schroeder, S. 2006b; Heal 1994). This will be discussed in detail in the next chapter.

The thrust of the argument is that many of the features that normally hang together with the concept 'description' (observation, object, state, etc.) are absent from our application of the term to psychological self-ascriptions, and that to call such utterances 'descriptions' is therefore liable to mislead. The point is made explicitly in a number of other places:

56 Wittgenstein claims that ""I intend…" is *never* a description". This, however, seems to be a touch hyperbolic. Although it isn't a description of a state of mind, it *can* be used to preface a description of *what one intends to do*. Someone may request me to 'Describe what you intend to do' and I may say 'I intend to…'.

57 Although it is a disposition of a special kind, unlike, say an inclination towards jealousy, which one can learn about through self-observation (cf. RPP II §178).

To call [“I have a toothache”] a description is misleading in a discussion like this.

What we call the description of a feeling is as different from the description of an object as “the name of a feeling” is different from “the name of an object”. (LSD, p.11)

To call the expression of a sensation a *statement* is misleading because 'testing', 'justification', 'confirmation', 'reinforcement' of the statement are connected with the word “statement” in the language-game. (Z §549)

For one or another of a class of features goes with what we call “describing”. Observing, considering, remembering behaviour, a striving for accuracy, the ability to correct oneself, comparing. (LW I §51)

Wittgenstein, however, was not unaware of the fact that there clearly *are* differences between an inarticulate cry and, for example, the words 'I am in pain' or 'I am afraid', even when both are uttered as a spontaneous manifestation of pain or fear:

A cry is not a description. But there are transitions. And the words “I am afraid” may approximate more, or less, to being a cry. They may come quite close to this and also be far removed from it…

We surely do not always say that someone is *complaining*, because he says he is in pain. So the words “I am in pain” may be a cry of complaint, and may be something else.

> But if “I am afraid” is not always something like a cry of complaint and yet sometimes is, then why should it *always* be a description of a state of mind? (PI, p. 189)

Here, Wittgenstein appears to concede that 'I am afraid' is *sometimes* a description of a state of mind. Indeed, earlier in PI II ix, Wittgenstein asks the question directly:

> Are the words “I am afraid” a description of a state of mind? (PI, p. 187)

He does not answer the question here, but in a previous draft, he went on to answer the question neither 'yes' as the descriptivist has it, nor 'no' as the expressivist has it, but rather, as Macarthur rightly points out:

> It depends on the game they are in. (LW I §412)

The words 'I am afraid' may be uttered in all sorts of contexts: A cry of fear is wrung from us when we are afraid – it is an involuntary response. However, I may, for example, cry out 'I'm in pain!' with the intention of stopping you from doing something that's causing me pain, in which case the utterance is not simply a cry of pain; I cried out because I wanted you to know that I'm in pain so that (hopefully, given that you care about the fact that I am in pain) you would stop what you were doing. Even further from a cry, perhaps, is a 'confession'. If a cry is 'wrung from us' involuntarily, then a confession, which it may take a great effort of will to utter, is no cry (PI §546). Yet none of these is what we would ordinarily call a 'description'.

Again, as Wittgenstein notes

> “I'm afraid” can, for instance, be said just as an explanation of the way I'm behaving. In that case it's far from being a groan; it can even be said with a smile. (LW I §21)

Wittgenstein, however, did allow that many psychological self-ascriptions could legitimately be called 'descriptions':

> If I tell someone “I can't keep my mind on my work today; I keep on thinking of his coming”— *this* will be called a description of my state of mind. (PI §585)

> “I am revolving the decision to go away to-morrow.” (This may be called a description of a state of mind.)… (PI §588)

Even in cases where Wittgenstein is happy to call a psychological avowal a description, however, this concession nonetheless comes with a caveat:

> “What is fear?” – “Well, the *manifestations* and occasions of fear are as follows: - - - ” – “What does 'to be afraid' mean?” – “The expression 'to be afraid' is used in *this* way: - - -”.
>
> “Is 'I am afraid - - -' therefore a description of my state?” It can be used in such a connection and with such an intention…But if, for example, I simply want to tell someone about my apprehension, then it is not that kind of description. (LW I §20)

Here, Wittgenstein draws a distinction between 'kinds' of description and it appears that the distinction he wished to draw was that between, on the one hand, descriptions of objective states of affairs in the world, such as the description of the *manifestations* of fear, or of the use of the word 'fear' in ordinary language, where mistakes are possible, and, on the other hand, carefully considered characterizations of one's own fear, where only indeterminacy and indecision, but not doubt and error, are possible. Someone may ask me to describe how I am feeling, and I may, quite legitimately reply 'I am afraid', and it will be said that I described myself as being afraid, but there is no possibility of doubt or error here. Similarly, my doctor may well ask me to describe my pain and I may say 'I have a dull ache in my left leg'. This will certainly be called a 'description' of my pain, but it is not based on observation of my behaviour, nor yet of my pain. There is, again, no possibility of genuine doubt or error here, and no observation is involved. What I say may, in a way, be wrong, because, for example, of a failure of presupposition (I may have no left leg) (Hacker 1990a, p. 21), but this is not to have misidentified my pain – I am not wrong about the fact that I *feel* a pain in my left leg, or at least, where my left leg *was* (Hacker 1990a, p. 21). I may also subsequently decide that my initial characterisation of my pain was inadequate in various respects, but again, this will not have been because I failed properly to observe my pain, but because I failed to be sufficiently articulate or to choose the most appropriate words to express my pain (Hacker 1990a, pp. 94f).

As previously noted, however, Wittgenstein thought that the negative affinities of such utterances with descriptions of objects justified a difference in terminology, at least in the context of a philosophical discussion. Indeed, for the avoidance of confusion, Wittgenstein generally preferred to reserve the term 'description', for cases, such as the description of an object, where doubt and error are possible. Wittgenstein tended, instead, to use the term 'report' for cases in which an avowal is issued coolly, with the intention of

conveying information, but where its truth nonetheless coincides with its truthfulness (RPP I §600).

In fact, Wittgenstein went as far as suggesting that the differences between cases of expressing one's fear could justify the use of different words for 'fear' in each case:

> I say “I am afraid”; someone asks me: “What was that? A cry of fear; or do you want to tell me how you feel; or is it a reflection on your present state?” – Could I always give him a clear answer? – Could I never give him one?
>
> We can imagine all sorts of thing here, for example:
> “No, no! I am afraid!”
> “I am afraid. I am sorry to have to confess it.”
> “I am still afraid, but no longer as much as before.”
> “At bottom I am still afraid, though I won't confess it to myself.”
> “I torment myself with all sorts of fears.”
> “Now, just when I should be fearless, I am afraid!”
> To each of these sentences a special tone of voice is appropriate, and a different context.
>
> It would be possible to imagine people who as it were thought much more definitely than we, and used different words where we use only one. (PI, pp. 187f)

Although we are not such beings, we could imagine people who consciously considered each utterance before it was made and thus used different words to signify their intent on

that occasion.

The same form of words may be *meant*, i.e. *intended* as a complaint, a confession, or a description and the difference between these cases need not reside in anything that accompanied the utterance which is identifiable as one's 'intention'. What one *meant* by an utterance is determined by the context of the utterance, including what one sincerely *says* one meant, when asked.

Nothing need have accompanied my words, as a sensation of pain accompanies an expression of pain, in order for me to have intended them as a report, rather than as a cry of complaint. I needn't have consciously *thought* about what I meant by an utterance when I said it, but it doesn't follow that I didn't mean anything by what I said:

> And do I always talk with very definite purpose?—And is what I say meaningless because I don't? (PI, p.188g)

That my utterance was a description is determined by the circumstances of the utterance, the manner in which it is uttered and what I say I meant when asked. Thus it is no great surprise that the same words are used in very different ways:

> Describing my state of mind (of fear, say) is something I do in a particular context. (Just as it takes a particular context to make a certain action into an experiment.)
>
> Is it, then, so surprising that I use the same expression in different games? And sometimes as it were between the games? (PI, p. 188)

Indeed, it needn't always be possible to specify exactly what one's intentions were in making an utterance: to express, report or describe one's psychological state:

> I say “I am afraid”; someone asks me: “What was that? A cry of fear; or do you want to tell me how you feel; or is it a reflection on your present state?” – Could I always give him a clear answer? – Could I never give him one? (PI, p. 187)

Wittgenstein does not answer these questions directly in PI II ix. Elsewhere, however, he makes it clear that the answer to both questions is 'no':

> I say “I am afraid...”, someone else asks me “What did you want to say when you said that? Was it like an exclamation; or were you alluding to your state within the past few hours; did you simply want to tell me something?” Can I always give him a clear answer? Can I never give him one? – Sometimes I shall have to say: “I was thinking about how I spent the day today and I shook my head, vexed with myself, as it were” – but other times:
>
> “It meant: Oh, God! If I just weren't so afraid!” – Or: “It was just a cry of fear” – or: “I wanted you to know how I feel.” Sometimes the utterance is really followed by such explanations. But they can't *always* be given. (LW I §17)

Sometimes, if asked what the purpose of an utterance was, we can give an answer, such as 'I wanted to tell you how I was feeling', or 'It was just a cry of fear', we might even imagine

different words being used in order to signify that an utterance was meant as a report of fear, say, rather than as a cry of fear. On other occasions, however, such explanations fail us and the only reply we can give is to paraphrase, or elaborate upon, what was said:

> If I tell you "I have been afraid of his arrival all day long" – I could, after all, go into detail: Immediately upon awakening I thought.... Then I considered.... Time and again I looked out of the window, etc., etc. This could be called a report about fear. But if I then said to somebody, "I am afraid..." – would that be as it were a groan of fear, or an observation about my condition? – It could be either one, or the other: It might simply be a groan of fear; but I might also want to report to someone else how I have been spending the day. And if I were now to say to him: "I have spent the whole day in fear (here details might be added) and now too I am full of anxiety" – what are we to say about this mixture of report and statement? Well what should we say other than that here we have the use of the word "fear" in front of us? (RPP II §156)

What is clear is that Wittgenstein was very much alive to the nuances of psychological language. Whilst the most primitive of such expressions are more or less just verbal expressions 'grafted on' to natural expressions, we soon outstrip such a limited use of these words. Expressions of pain, fear and the like can be wrung from us, like a cry, or torn from us, like a confession. They may be blurted out in a panic, or issued coolly after careful consideration, in order to inform. All of these uses fall under the same psychological concepts, but we should be careful to distinguish between them, lest philosophical confusion ensue.

Generally, what we ordinarily call 'descriptions' of our psychological states have in common one or more of a number of features: their cool delivery, their informative intention, their level of detail, and so on. But the criteria for what constitutes a 'description' of one's own psychology are far from rigid.

Moreover, whilst usually unobjectionable, in a philosophical setting, calling such utterances 'descriptions' can be misleading. For 'descriptions' usually involve observation of an object, testing, the possibility of doubt, error and correction, whereas none of these applies in the case of one's own pain, fear, and the like.

None of this, however, impugns the notion that such utterances are *also expressions* of one's psychological states. Wittgenstein never denied that what can legitimately be called a description of a psychological state can also be an expression of that state.

Ultimately, the moral of Wittgenstein's discussion of psychological self-descriptions is that one should not ask such a shape-shifting concept as 'description' to carry too much philosophical weight. It needn't indicate the existence of any object of description, nor that any kind of observation is involved. Hence, whilst Macarthur is correct to point out that Wittgenstein did not wish to banish the term from our everyday psychological discourse, he warned heavily against the philosopher employing it without due care and attention.

'The Problem'

Wittgenstein, however, saw that there was a problem even with the notion that primitive psychological avowals are not descriptions:

> But here is the problem: a cry, which cannot be called a description, which is more primitive than any description, for all that serves as a description of the inner life. (PI, p.189b)

That 'I am in pain' is a grammatically declarative English sentence leads to a puzzle, for it engenders a seeming symmetry between first and third person psychological utterances, and between first-person psychological utterances and statements about objects. Both 'I am in pain' said by me and 'He is in pain' said about me by someone else are true if and only if I am in pain. From the truth of his avowal 'I am in pain', it follows that he is in pain, just as from the truth of his assertion 'My room is beige', it follows that his room is beige (RPP II §728; LW §43). Furthermore, first-person present avowals can be used in the same kinds of logical inferences as can their third-person counterparts. Thus the conclusion becomes almost irresistible that such first-person avowals say of oneself just what their third-person counterparts do, and, in a way, of course, they do, for there is a perfectly trivial sense in which both 'I am in pain' said by me and 'He is in pain' said about me by someone else say that I am in pain and are made true by the fact that I am in pain. Thus, 'I am in pain', even when uttered as a substitute for a primitive expression of pain, which nobody but a philosopher would call a description, appears to function as a description after the manner of 'He is in pain'.

Whilst this misconception is dealt with most thoroughly in Wittgenstein's treatment of Moore's Pradox, which we shall discuss in some detail in the next chapter, Wittgenstein was also keen to combat the notion that psychological self-ascriptions constitute descriptions simply because the same thing can be inferred from the truth of both first and third person present tense psychological ascriptions, namely that someone is in a particular 'state of mind'. Thus, Wittgenstein says:

> It might be said: An assertion says something about the state of mind, given that I can make inferences from it [to] the state of mind. (That sounds more stupid than it is.) If *that's* how it is, then the expression of a wish "Give me that apple" says something about my state of mind. And is this proposition then a description of this state? That one won't want to say. ("Off with his head!") (RPP I §463)

And similarly:

> The sentence "I want some wine to drink" has roughly the same sense as "Wine over here!" No one will call *that* a description; but I can gather from it that the one who says it is keen to drink wine, that at any moment he may take action if his wish is refused – and this will be called a conclusion as to his state of mind. (RPP I §469)

As Wittgenstein points out, a 'state of mind' i.e. a wish, can be inferred no less from an imperative than from a psychological ascription, but nobody will call *these* 'descriptions'.[58]

Indeed, one can also infer a 'state of mind' from natural expressive behaviour in a

58 It may be thought that the imperatives which Wittgenstein here takes to be more or less synonymous with sentences self-ascribing a wish or desire are not so. In fact, 'I want that apple' does not even follow from 'Give me that apple', even where this order is given in all seriousness, for one may issue an order which one does not wish to be carried out, as when an army officer, whilst disagreeing with his commander's aims, nonetheless issues an order out of duty.

It should be noted that, even if this objection worked, its consequences would be somewhat limited, as it could be said that the commander is here speaking through his subordinate's mouth, such that the commander's and not the speaker's wishes are expressed, thus the function of an order is to express *someone's* wishes, but still not to describe them.

However, 'Wish' needn't mean 'desire', where this is taken to be some strongly felt emotional state. I may wish an order to be carried out, even though I give it with a heavy heart, as when I choose between the 'lesser of two evils'.

Furthermore, if an officer does not take steps to ensure that his orders are carried out, then it would seem that his order was somewhat less than a genuine one. In one sense, he has performed a speech-act of giving an order, but it is not a 'happy' one (the order has been issued with something akin to insincerity.) Yet, if someone takes all necessary steps to ensure that an order is carried out, then the assertion that the person giving the order did not 'wish' it to be carried out becomes suspect, even if he didn't do it happily.

certain context:

> Can I not say: a cry, a laugh, are full of meaning?
>
> And that means, roughly: much can be gathered from them. (PI §543)

But such behaviour will certainly not be called a 'description of a state of mind'. I may, after all, be taught to cry out 'Ouch!', 'It Hurts!' or 'I am in pain!', instead of simply issuing an inarticulate cry, but, as Wittgenstein notes elsewhere:

> If a cry is not a description, then neither is the verbal expression that replaces it. (RPP II §728; cf. PI, p. 187d-e)

For it would seem that, in the context in which I simply cry out in response to pain, the difference between an inarticulate cry and the verbal expression that replaces it is simply that I have been trained to make a different sound, and this can hardly turn what was not a description into a description.

Observing, Describing *and* Expressing

So far we have discussed first-person present psychological avowals, constituting expressions of 'psychological states'. These utterances are characterised by the fact that their truth follows from their truthfulness (PI p. 226). Such statements are not based on observation and are thus not susceptible of genuine doubt or error, but only indeterminacy or indecision.

Unlike first-person present psychological utterances, statements about one's past and future psychological states *can* be said to be based on observation (PI, p. 187b-c). It is possible for one to doubt and to be mistaken about whether one *was* in pain at a particular time, to make sure, find out, and thus to *know* that one was in pain, just as it is possible to find out and thus to know that another person is or was in pain. Such utterances are obviously not expressions of pain, but expressions of belief (cf. Z §79; RPP II §4), as where doubt and error are possible, so too are belief and knowledge (cf. RPP I §142; RPP II §§169, 282; PI, p. 192b; Hacker 1990a, p. 31).

My knowledge that I *was* in pain, however, does not usually have the same grounds as that of other people, e.g. my avowals and behaviour at the time and my subsequent claims to have been in pain (RPP I §§466ff). Of course, one *can* come to know that one was in pain at some time by, say, observing a video of oneself after an accident of which one has no memory, or from the testimony of others who witnessed the accident. In the normal case, however, to *know* that one was in pain is simply not to have forgotten that one was in pain. (Hacker 1990a, p. 33) Assuming there is no reason to doubt the reliability of a person's memory, it will usually be granted that he *knows* whether or not he was in pain at a particular time.[59]

Even where a person's claim to have been in pain is somewhat at odds with his behaviour at the time which he is recalling, his sincere avowal will usually be accepted if he can provide sound reasons why this was the case. Thus if you tell me that yesterday you were in pain, but I object that yesterday you were up and active playing football with your children, you may say that you were 'putting a brave face on' for the sake of the children and I will not, therefore, doubt your memory (although I may still doubt your sincerity.)

59 Memory is closely analogous to sense-perception in that it can be a more or less reliable informant of the (past) facts. My memory can be tested in much the same way that my eyes can be tested (not, of course, by the same methods, but by asking me to recall events that are especially well-documented, or have been witnessed by others.) Of course, even if my memory, like my eyesight, is tested and found to be good, there remains the theoretical possibility that I have misremembered or misperceived in a particular case.

If, of course, your memory of yesterday's events seems faulty – for example you claim distinctly to remember that you were limping heavily on your right leg and wincing, when in fact you were doing no such thing – then your word may come in question. On the other hand, if your memories of your behaviour and its circumstances appear sound, yet your avowal to the effect that you were in pain conflict to a great extent with these memories – for example, you claim to have been in abject agony all day whilst admitting that you showed no sign of being so – then you will likely be accused of insincerity, or at least exaggeration, and if you insist that your failure to manifest pain in any way betokens either of these, your understanding of the word 'agony' will come in question. However, if your memories of yesterday's events are sound and you are able to provide reasons for discrepancies between your claim and your behaviour your word regarding whether or not you were in pain will be taken as authoritative, that is, it will be accepted that you remember that you were in pain. In fact, so long as a person's memory is *in general* sound, then even if he cannot remember much detail about an event, his sincere claim to have been in pain will usually be accepted, so long as it does not conflict with his behaviour or avowals at the time.

It remains *possible* that one is wrong about one's having been in pain. Whether one was so is a matter of what one *could* have sincerely said at the time and there is room for a discrepancy between this and what one subsequently says. However, the theoretical possibility of error is a feature of *all* genuine cases of knowledge.[60] The fact remains that in certain circumstances we will grant that a person *knows* that he was in pain at a certain time.[61]

In the normal case of remembering one's pain, claims about one's past pain may be said to be based on 'self observation', inasmuch as one can put oneself in a better position

60 As Wittgenstein says: “For 'I know' seems to describe a state of affairs which guarantees what is known, guarantees it as a fact. One always forgets the expression 'I thought I knew'” (OC §12; cf. OC §21).

to allow one's memories to come through clearly, by, for example, avoiding external distractions whilst considering the matter, by having one's memory 'jogged' by other people's accounts of the surrounding events or by photographs, and by ensuring that there are no factors which are distorting one's memory. However, to say that such claims are based on 'self-observation' may be misleading inasmuch as they are not based on observation of any object, e.g. of one's behaviour, etc. as are other people's claims about one's past pain (RPP I §§466ff).

There are cases, however, in which one may express one's current 'psychological state' by means of an utterance which is, at least in part, the result of observation and about which one can thus be doubtful or in error. Thus, in PI II ix, Wittgenstein says:

> A touch which was still painful yesterday is no longer so today.
>
> Today I feel the pain only when I think about it. (That is: in certain circumstances.)
>
> My grief is no longer the same; a memory which was still unbearable to me a year ago is now no longer so.
>
> That is a result of observation. (PI, p. 187)

Here, Wittgenstein gives examples of genuine descriptions based on observation, which might be false, even if truthful, but which are *also* expressions of one's pain and grief (or lack thereof,) as their truthfulness guarantees a particular 'psychological state' on the part

61 Of course, one can *imagine* all sorts of fanciful scenarios in which one's brain is somehow being manipulated in order to implant false memories, but a genuine doubt, no less than knowledge, requires grounds (OC §§4, 122, 288, 322, 323, 458, 516, 606) and the theoretical possibility of alternative explanations is not such a ground, but simply a reminder of the grammatical fact that *all* empirical propositions are subject to the possibility of error (Schroeder, S. 2006a, p.210). Even if we were to hedge every assertion with a caveat such as 'perhaps' (OC §450) or 'unless I am mistaken' (cf. OC §626), in order to recognise the logical possibility that things might turn out otherwise, this would not betoken what we call *doubt*, but would simply amount to an idiosyncratic way of expressing what we call certainty (cf. OC §120). That such hedges would have to be added to *every* assertion shows that they would not alter the language-game, and thus would be senseless (OC §627).

of the speaker.

It might be argued that Wittgenstein's examples constitute conjunctive utterances of the form 'Yesterday I felt pain; today I do not feel pain'. In such a case, whilst one may be in error about the first conjunct, it is *not* an expression of pain, whereas the second conjunct is a straightforward expression of the absence of pain and *not* a description based on observation. In response, it may be argued that this still means that the conjunction as a whole is something about which one can be in error, thus qualifying it as a genuine description based on observation, whilst at the same time being an expression of the absence of pain.

However, we can imagine other examples which do not seem susceptible to such objections. Take, for example, the claim 'My pain is less severe than it was yesterday'.[62] Such an utterance cannot similarly be explained away as a conjunction of a description of one's past pain and an expression of one's current pain, for whilst such an utterance certainly implies 'I am in pain', it seems mysterious what the second conjunct could be other than 'My pain was worse yesterday', which is equivalent to the original assertion.

It is quite possible that one has misremembered the relative severity of one's pain on the relevant days. One may, therefore, doubt or be unsure in this regard and one may be legitimately challenged regarding the truth of such an assertion by someone else, even on the assumption that the utterance is truthful. Such an utterance, then, qualifies as a genuine description based on observation (not of one's behaviour, nor of an 'inner object', but of the changes in one's pain over time) but at the same time, its sincerity entails that the speaker is in pain and thus also qualifies as an expression of pain.

Similarly, a prediction such as 'I'll be in less pain by morning', is an expression of one's current pain, but one can doubt that it is the case or be wrong about its being so, and

62 Wittgenstein gives some similar examples at RPP II §728: “The utterances of fear, hope, wish, are not descriptions; but the sentences “I'm less afraid of him now than before”, “For a long time I've been wishing ...”, ... are descriptions.”

one can also *know* that it is true by experience (one regularly has these kinds of cramps and remembers that they always pass by the next day, or the doctor has told one so (and he has seen many cases just like this.))

It may be objected that in asserting a proposition such as 'My pain is less severe than it was yesterday' one merely 'conversationally implicates' that one is in pain, since one could always add 'In fact, my pain is gone, so it's *much* less severe', thereby cancelling the implicature.

To do this, however, would not be to cancel an implicature, but to withdraw one's original statement as inappropriate (such an answer may, of course, be given by way of a joke.) To answer the question 'Is your pain less severe than yesterday?' by saying 'Yes, it's gone completely' is like answering the question 'Have you stopped beating your wife?', by saying 'No, I never beat her in the first place'. The question 'Is your pain less severe than yesterday?' presupposes that one is *still* in pain. One might, on being asked such a question, say 'Actually, it's gone', which would be to reject the presupposition of the question (RPP I §274).

A Model for Ethics?

Suppose that we had just one word which was used in a case of simple predication to mean what is meant by 'My pain is less severe than it was yesterday'. Perhaps, we could introduce the convention that 'improving' was to serve this purpose, such that, for example, the doctor asks me 'How are you today?' and I say 'I am improving'. 'I am improving' would be an expression of pain, but it would also be possible for me to be wrong about the fact that I am improving, for my memory may be unsound, and I may have forgotten that

yesterday my pain was better controlled by medication.

'I am improving', though, needn't be used only in the context of a considered response to my doctor's question. I might also awake and exclaim with surprise or delight 'I am improving!', with no intention of telling anybody anything, or I may even say this to myself, by way of self-encouragement (cf. PI §585).

The use of such a word, I believe, would be somewhat similar to the use of a word like 'Wrong' in its ethical sense. Whilst '*x* is wrong' is an expression of an attitude[63] of disapproval towards *x* it is also a statement about which I can be in error, or the truth of which I can often doubt. Furthermore, whilst '*x* is wrong' can be said as an exclamation, akin to a sigh of disappointment, when, for instance, I see something appalling on the news and exclaim 'That's just wrong!', it can *also* be used coolly with the purpose of characterising or describing an action or form of behaviour in terms of my attitude towards it, or with the intention of letting someone know what my attitude towards a thing is.

The difficult question, of course, is what constitutes a standard of correctness for my application of the term 'wrong'. This, it seems, can be answered by considering how we do in fact decide whether something is right or wrong, good or bad, and the answer to that question, it seems, is (broadly) that we consider the features of the thing in question and our responses to them, noting, amongst other things, the consistency or inconsistency of our responses and making decisions, in cases of conflict, as to which of our responses is to be upheld and which discarded. Of course, not just *any* consideration will do. I must make no mistakes in reasoning and my immediate responses ought not to be clouded by selfishness, anger and the like.

If moral terms have such a function, then, whilst they can be used in spontaneous expressions of moral indignation that won't be called descriptions (either of my attitude or of a form of behaviour) they can *also* be used in order to describe actions and forms of

63 Which is *not* a 'mental state', but more akin to a disposition.

behaviour in terms of one's reactions to them, whilst simultaneously serving to express one's attitudes. I might, for example, describe my attitude towards an action by saying 'I would disapprove of this after careful consideration', or I might equally put this in terms of a description of an object by saying 'This is such that, after careful consideration, I would disapprove of it.' In this way, moral judgements can even be based on 'observation', given that one can put oneself in a better position to be sure of what one says by considering the matter carefully. Moreover, someone who is aware of my moral standards will be able to garner something of a description of the action's features from what I say. It is now the task of the remainder of this thesis to attempt to outline a metaethic which allows moral terms to have such a use.

IV

Moore's Paradox and the Parity Thesis

As I argued in Chapter Two, Wittgenstein's remarks on ethics have distinct affinities with contemporary non-cognitivism or expressivism. We ought, however, not to rush into following Wittgenstein in this regard, for, as I shall now begin to show, whilst such a view captures many of the features of moral language we are interested in, it also brings up philosophical perplexities that can, perhaps, be avoided by taking a slightly different tack.

Metaethical non-cognitivism is often defined in opposition to simple subjectivism as the thesis that a moral utterance such as '*x* is wrong' serves not to *report*, but to *express* one's attitude of disapproval towards *x*.[64] In invoking a distinction between reporting and expressing an attitude, non-cognitivists hope to retain simple subjectivism's central virtue of accounting for the conceptual connection between moral judgements and attitudes such as approval and disapproval, whilst overcoming its chief vices: that it renders moral judgements objectionably mind-dependent and cannot account for moral disagreement. A number of commentators, however, have sought to show that non-cognitivism collapses into simple subjectivism and therefore fails to overcome these faults.

In this chapter, I shall explain that such arguments falter in failing to respect the contemporary non-cognitivist's 'Parity Thesis'. The discussion, however, will reveal that the relationship which, according to the Parity Thesis, holds between self-ascriptions and expressions of belief and attitude is captured neatly in 'Moore's Paradox'.

I shall argue that Wittgenstein's solution to Moore's Paradox reveals that the

64 This is, perhaps, not an entirely accurate definition of non-cognitivism as a whole, but instead a definition of 'expressivism', a particular form of non-cognitivism. However, emotivists, too, accept something like this claim (cf. Ayer 1936, p.110) and even prescriptivists who claim that moral terms are used to issue prescriptions, generally hold that prescriptive language is also expressive of wishes, attitudes, etc. (cf. Hare 1952, p. 10).

relationship between '*p*' and 'I believe *p*' is that, as *utterances*, they usually have the same meaning, whereas, as *sentences* they differ in meaning. Similarly, as utterances, 'I disapprove of *x*' and '*x* is wrong' usually have the same meaning, even though as sentences they are not synonymous.

I shall conclude that, whilst non-cognitivism may represent a correct account of the standard meaning of an utterance of an atomic moral sentence, it fails to provide a *complete* account of the meaning of the sentence itself and of the moral terms therein contained. In fact, as I shall argue, no such account *can* follow from the non-cognitivist's explanations of the meanings of moral utterances. Thus, the non-cognitivist still has work to do in responding to the Frege-Geach problem.

Expressivism and the Parity Thesis

A central claim of contemporary metaethical non-cognitivism (or 'expressivism') can be summed up in a schema dubbed the 'Parity Thesis' (Schroeder, M. 2010a, p. 77; Schroeder, M. 2010b, p. 3):

Parity Thesis: *x* is wrong : I disapprove of *x* :: *p* : I believe that *p*[65]

The Parity Thesis says that '*x* is wrong' stands to 'I disapprove of *x*' as '*p*' stands to 'I believe that *p*' and is perhaps best understood as a response to some well-worn problems with 'simple subjectivism': the claim that '*x* is wrong' means 'I disapprove of *x*'.

The 'disagreement problem' is, roughly, the claim that simple subjectivism is false,

65 For simplicity, I shall talk throughout about the meaning of the word 'wrong' and of sentences and utterances containing this word. However, what I say can be applied *mutatis mutandis* to other moral terms.

because when one person says '*x* is wrong' and another says '*x* is not wrong' their claims are inconsistent, whereas when one person says 'I disapprove of *x*' and another 'I don't disapprove of *x*' their claims are consistent.

The 'modal problem' is, roughly, the claim that simple subjectivism is false because it entails that '*x* is wrong iff I disapprove of *x*' is an analytic truth, whereas it is, in fact, patently false (unless, perhaps, God says it).

The expressivist notes that a non-moral subjectivism, according to which '*p*' meant 'I believe that *p*', would encounter similar difficulties. For if one person says '*p*' and the other 'not-*p*' they contradict one another, whereas if one says 'I believe that *p*' and the other 'I don't believe that *p*', they apparently do not. Similarly, on such an account, '*p* iff I believe that *p*' ought to be a theorem, whereas it is, in fact, patently false (unless, perhaps, God says it).

The expressivist, therefore, simply claims that, the relationship, whatever it may be, which holds between '*p*' and 'I believe that *p*', also holds between '*x* is wrong' and 'I disapprove of *x*'. Since this relationship, whatever it is, does not give rise to a disagreement or modal problem in the non-moral case, there cannot, given the Parity Thesis, be such a problem in the moral case either (Schroeder, M. 2010b, p. 2).

Avoiding Collapse: Three Examples

As recently pointed out by Mark Schroeder (2010b & 2014), the Parity Thesis dispenses neatly with a number of arguments which claim to show that expressivism collapses into simple subjectivism:

1. Jackson & Pettit: Lockean and Logical Truth-Conditions

Frank Jackson and Phillip Pettit (1998) argue, following Locke, that the meaningfulness of a word depends upon a convention amongst speakers to use it only under certain conditions. Such 'conditions of correct use', claim Jackson & Pettit, just are truth-conditions (Jackson & Pettit 1998, p. 241). Jackson & Pettit hold that, on an expressivist account, the conditions of correct use of a moral term are that the speaker has the attitude which it is used to express and thus conclude that a speaker's having the appropriate attitude constitutes the truth-condition of his moral judgement. Thus non-cognitivism collapses into subjectivism.[66]

As has now been pointed out by a number of commentators, however, the application of Jackson & Pettit's argument to the non-moral case in accordance with the Parity Thesis yields the implausible conclusion that the truth-condition of a non-moral assertion is that the speaker believes what he asserts (Dreier 2004; Ridge 2006, p. 506).[67]

Jackson & Pettit conflate truth-conditions with what John Searle called 'sincerity conditions' (Searle 1970, p. 65; Ridge 2006, p. 507).[68] They thus convict expressivism of collapse into simple subjectivism only at the expense of rendering *all* assertions objectionably mind-dependent.

Of course, there *are* cases where one cannot speak sincerely, yet falsely, i.e. self-ascriptions of belief, attitude, etc. That moral assertions are such self-ascriptions, however,

66 In fact, according to Jackson & Pettit, words essentially stand for objects and our ability to use a word correctly rests on the possibility of *recognising* the object for which it stands. Thus, a term which serves to express an attitude is applied correctly only when the speaker *believes* that he has that attitude. I take the notion that one might *believe* that one has an attitude (in the relevant sense) to be incoherent and so address what I take to be a more plausible form of Jackson & Pettit's argument. However, their argument fails even on its own, less plausible, terms (see Ridge 2006, pp. 504-7).

67 Indeed, Jackson & Pettit (1998, p. 241) appear to endorse this conclusion explicitly, claiming e.g. that it is correct to apply the word 'square' only if one *takes a thing* to be square.

68 Jackson & Pettit are in fact quite explicit on this point: “expressivists must allow that we use the word sincerely only when we *believe* that we have a certain kind of attitude. And then it is hard to see how they can avoid conceding truth conditions to 'That is good', namely, those of that belief” (Jackson & Pettit 1998, p. 242).

is precisely what expressivists deny. According to the Parity Thesis, expressivists are no more committed to this claim than to the claim that non-moral assertions are self-ascriptions of belief.

2. Shafer-Landau: Biconditionals

Like Jackson & Pettit, Russ Shafer-Landau (2003, pp. 30-33) argues that, on an expressivist account, there can be no standard of correctness for the application of a moral term other than the actual moral outlook of a speaker. Thus, all speakers are committed to the truth of biconditionals of the form '*x* is wrong iff I disapprove of *x*'.[69]

The idea here seems to be that, on an expressivist account, it would be incoherent to assert either

(1) *x* is wrong, but I don't disapprove of *x*.

or

(2) *x* is not wrong, but I disapprove of *x*.

Thus it appears that speakers are committed to the negations of these statements, which are logically equivalent to

69 It seems to me that Shafer-Landau's argument may in fact be more subtle than this. In fact, in a more recent version of Schroeder's paper (Schroeder, M. 2014), he omits reference to Shafer-Landau, perhaps for this reason. However, Schroeder appears not to be alone in interpreting Shafer-Landau in this way (see Suikkanen 2009, p. 366: fn12) so it seems worth mentioning this reading, regardless.

(1') If *x* is wrong, then I disapprove of *x*.

and

(2') If I disapprove of *x*, then *x* is wrong.

respectively.

If it is incoherent to deny both (1') and (2'), however, then it is surely incoherent to deny their conjunction, which is logically equivalent to

(3) *x* is wrong iff I disapprove of *x*.

Thus expressivism, once again, collapses into simple subjectivism.

The application of Shafer-Landau's argument to the non-moral case in accordance with the Parity Thesis, however, once more provides a *reductio*. For it is also incoherent to assert either

(4) I believe that *p*, but not-*p*.

or

(5) I don't believe that *p*, but *p*.

But nobody would argue that speakers are therefore committed to the patently absurd claim

(6) *p* iff I believe that *p.*

The observation that it would be incoherent to assert (4) or (5) is known as 'Moore's Paradox' and Shafer-Landau has succeeded only in showing that, on a non-cognitivist account, Moore's Paradox arises in the moral sphere, just as it does in the non-moral one.

But this is a problem for *everyone*, not just expressivists. After all, expressivists accept the equivalence claim

(7) I believe that *x* is wrong iff I disapprove of *x.*

For expressivists, then, (3) is equivalent to

(8) *x* is wrong iff I believe that *x* is wrong.

But that it seems, paradoxically, to be incoherent to deny the obviously false (8) is just as much a problem for cognitivists as for non-cognitivists, so Shafer-Landau hasn't shown that expressivists have any special problems of their own to overcome (Schroeder, M. 2010b, pp. 3-6).

3. Suikkanen: 'Positive Semantic Evaluation'

Jussi Suikkanen (2009) argues that sentences have 'conditions of positive semantic evaluation' such that every competent speaker must accept

(P1) *x* is wrong iff I can now use the sentence '*x* is wrong' to make a statement that receives a positive semantic evaluation.

According to Suikkanen, expressivism entails

(P2) I can now use the sentence '*x* is wrong' to make a statement that receives a positive semantic evaluation iff I now disapprove of *x*.

These premises jointly entail

(C) *x* is wrong iff I now disapprove of *x.*

Thus, Suikkanen concludes, expressivism collapses into simple subjectivism (Suikkanen 2009, pp. 368-75).

If, however, as Suikkanen (2009, p. 370) suggests, 'receives a positive semantic evaluation' means 'is sincere', then, whilst (P2) is indeed entailed by expressivism, (P1) certainly isn't. For, by parity of reasoning, we can construct its patently false non-moral equivalent

(P1') Grass is green[70] iff I can now use the sentence 'Grass is green' to make a statement that *is sincere.*

70 From a Wittgensteinian, perspective, this may not, perhaps, be the best proposition to use in such an example, as there may be complications surrounding the question whether or not this is a grammatical proposition. However, I shall ignore such complications here in order to stick with the examples Mark Schroeder uses.

If, on the other hand, 'receives a positive semantic evaluation' means 'is true', then (P1) will be true, but (P2) will be patently false, for, by parity of reasoning, we can construct

(P2') I can now use the sentence 'Grass is green' to make a statement that *is true* iff I now believe that grass is green.

Truth-conditions, once again, are not sincerity conditions (Schroeder, M. 2010b, pp. 6f).

In fact, as Mark Schroeder points out, the crux of the matter is that the denial of (C) would be logically equivalent to

(C') Either *x* is wrong and I don't disapprove of *x*, or I disapprove of *x* and *x* is not wrong.

and, of course, the expressivist takes it to be incoherent to assert either disjunct.

But this, once again, is simply Moore's Paradox, and precisely the same problem arises in the non-moral case. For the denial of

(6) *p* iff I believe that *p.*

is logically equivalent to

(9) Either *p*, but I don't believe that *p,* or not-*p*, but I believe that *p.*

and it would be as incoherent to assert either disjunct of (9) as it would be to assert either

disjunct of (C'). That the denial of the obviously false (6) engenders Moore's Paradox, however, is a problem for everyone and not just for expressivists (Schroeder, M. 2010b, pp. 8f). The argument again shows only that, on an expressivist account, Moore's Paradox arises for moral claims just as for non-moral ones and not that expressivism collapses into simple subjectivism.

For expressivists, '*x* is wrong' and 'I disapprove of *x*' no more have the same truth conditions than do '*p*' and 'I believe that *p*'. Any argument which identifies sincerity conditions with truth conditions in the moral case ignores the Parity Thesis and thus misses its target, or else has absurd consequences when the Parity Thesis is applied.

The Significance of Moore's Paradox

Expressivists nonetheless have questions to answer regarding their Parity Thesis. Most pressingly, they owe an account of just what the relationship between '*p*' and 'I believe that *p*' *is*, such that it can also hold between '*x* is wrong' and 'I disapprove of *x*' in a way that is consistent with their thesis that moral claims, unlike non-moral ones, do not serve to describe any feature of the world.[71]

From the foregoing, however, we can say *something* about the relationship which holds between '*x* is wrong' and 'I disapprove of *x*' as between '*p*' and 'I believe that *p*', viz. That, in both cases, it would be Moore-paradoxical to assert one whilst denying the other.

An explication of the roots of Moorean absurdity, then, would seem helpful in characterising the relationship identified by the Parity Thesis. Indeed, it would seem that

71 It seems to me that this may be the challenge Shafer-Landau really had in mind. A deep-seated puzzlement about just what truth-conditions (or conditions of correct use) might *be* on an expressivist account, if they are not sincerity conditions, seems, with some justification, to run through all of the above arguments.

no understanding of this relationship is possible apart from such an explication.[72]

The Paradox

It is undeniably possible that, whilst I don't believe that I went to the pictures last Tuesday, I nonetheless did so. I may, after all, have forgotten that I did. Similarly, it is undeniably possible that, although I believe that I had tea with breakfast this morning, I did not in fact do so. I may, after all, have misremembered.

It is, however, equally undeniable that to assert either

(10) I don't believe that I went to the pictures last Tuesday, but I did.

or

(11) I believe that I had tea with breakfast this morning, but I didn't.

or, more generally, to make assertions of the form

(4) I believe that *p*, but not-*p*.

or

(5) I don't believe that *p*, but *p*.[73]

72 It is interesting to note that Moore's Paradox first arose in the context of an argument about the distinction between subjectivism and non-cognitivism (see Moore 1942).

would usually[74] be incoherent. Were someone, in all seriousness, to make an assertion of this form we would not understand him; we would say that he had *contradicted* himself (MS 132, 119f).

This, then, is Moore's Paradox: whilst it is *possible* that I believe that *p*, when not-*p*, or that I don't believe that *p*, when *p* (and whilst I can recognise this possibility explicitly), it would nonetheless be (in some sense) contradictory for me to assert[75] that either of these possibilities obtains. Thus, in asserting something that is apparently not contradictory, I have contradicted myself, which appears to amount to a contradiction in terms!

Or again the assertions

(12) It's possible that I believe that *p*, but not-*p*.

and

(13) It's possible that I don't believe that *p*, but *p*.

make perfect sense, whereas the assertions (iv) and (v) are nonsense. How, then can it be nonsensical for me to assert something that is possible?[76]

73 NB – the paradox also arises (arguably in a different form (see Heal 1994)) for claims of the form 'I believe that not-*p*, but *p*'. I shall not address this complication here, but note with interest that, it corresponds to the so-called 'negation problem' for expressivists (cf. Unwin 1999 & 2001).

74 Of course, such utterances aren't *always* incoherent. One might say '*p*' to one person and 'I don't believe that *p*' to another, or one might say 'I don't believe that *p*', then immediately change one's mind and say 'But *p*!'. One may even have a split personality and believe that one's 'alter ego' was unaware that one had been to the pictures last Tuesday. Such cases, although bizarre, do not engender any genuine paradox. For here, 'I believe…' simply means 'My alter ego believes…' (cf. PI, p. 192).

75 Or even to *think* it (Heal 1994, p. 6; See footnote 8 below). I shall concentrate on the case of assertion, because I believe that an account of why it is impossible to have such thoughts ultimately rests on an account of why it is impossible to make such assertions. I cannot, however, go into detail on such matters here.

A Wittgensteinian Solution

It is highly natural to suppose that the utterances

(14) It's raining.

and

(15) I believe it's raining.

serve to describe two logically independent states of affairs: the state of the weather and one's state of belief, respectively. During his discussion of Moore's Paradox, however, Wittgenstein warns that:

> The difficulty becomes insurmountable if you think the sentence "I believe..." states something about the state of my mind. If it were so, then Moore's Paradox would have to be reproducible if, instead of saying something about the state of one's own mind, one were making some statement about the state of one's own brain. (RPP I §501)

76 It may be thought that in asserting '*p*', one 'conversationally implicates' that one believes that *p* and that this explains the paradox. Indeed, Moore himself, at one time, seems to have favoured an explanation along these lines (Moore 1942, p. 543). It is, however, a necessary condition (although not, as I argued in Chapter Two, a sufficient condition) of an implication's being the result of conversational implicature that it may be straightforwardly cancelled without absurdity (Grice 1989, pp. 39, 44), whereas the very examples in question are those in which one attempts to take back the purported implicature of one's assertion (i.e. that one believes what one asserts) with the next breath. Thus it is implausible to maintain that Moore's Paradox arises through implicature. Furthermore, such pragmatic considerations could only ever hope to account for the strangeness of *asserting* Moore-paradoxical sentences. It is, however, no less absurd to *think* (4) or (5) (Heal 1994, p.6) and the explanation as to why this is so cannot lie in any purely pragmatic feature of communication.

Were (15) a description of the speaker's state of mind on the model of a description of an object, then the utterances

(16) It's raining and I don't believe it's raining.

or

(17) I believe it's raining and it isn't raining.

would straightforwardly describe possible states of affairs (Schroeder, S. 2006b, p. 162) and Moore's Paradox would no more arise here than for statements such as

(18) It's raining and my brain is in state *S*.[77]

Wittgenstein, therefore, concludes that the natural picture one has of the use of 'I believe that *p*' – that of observing and subsequently describing one's 'inner state' of belief, just as one might observe the weather, or one's own brain, and subsequently describe these (RPP I §715) – is misleading. Despite appearances, 'I believe that *p*' is *not* a description of my state of mind.

It may appear, however, that one paradox has simply been replaced with another. For surely a person's believing something is a genuine state of affairs. I can describe *others* as believing something in such a way that no paradox arises, so why should this not be possible in my own case?

Wittgenstein's response begins with a suggestion about the nature of belief:

77 See Heal (1994) for further elaboration on this point.

> …Believing…is a kind of disposition of the believing person. This is shewn me in the case of someone else by his behaviour; and by his words. And under this head, by the expression “I believe...” as well as by the simple assertion… (PI, pp. 191f)

This suggestion might not seem to get us very far. My having a disposition is, after all, a state of affairs logically independent of the state of the weather, for example. So it seems it ought to be possible for me to describe myself as having the disposition called 'the belief that it's raining' whilst denying that it's raining. I ought to be able, on the one hand, to observe myself and to report my disposition, and, on the other hand, to observe and report on the state of the weather, quite independently.

As is implicit in the passage above, however, Wittgenstein's solution to this puzzle is that my avowal 'I believe that *p*' is a *manifestation* of the disposition which another person *describes* me as having when he says of me 'He believes that *p*' (cf. RPP II §281).

In many cases, one can self-ascribe a disposition without manifesting it. One can, for example, say of oneself that one has a tendency to become jealous (RPP II §178), to be quick to anger, or to be credulous (RPP I §502; Schroeder, S. 2006b, p. 168) without thereby manifesting these tendencies.[78] These kinds of disposition one self-ascribes, quite fallibly[79], on the basis of evidence, including memories of how one has behaved in the past. Avowals of belief, though, are manifestations of belief. They serve as criteria according to which others may ascribe to one a belief.[80]

Wittgenstein, though, anticipates a query:

78 It may, indeed, be misleading to call belief a 'disposition' if this difference between beliefs and other tendencies and capacities is thereby obscured (cf. RPP II §178).

79 Indeed, it is often *harder* to recognise such dispositions in oneself than it is to recognise them in others (Hacker 1990a, p. 31).

80 NB – It is not, for example, paradoxical to assert 'I'm gullible, but even I don't believe *that*' or 'Despite my irascible temperament, I'm not allowing him to anger me'.

> What about my own case: how do I myself recognize my own disposition? (PI, p. 192a)

His sardonic response to his foil is:

> Here it will have been necessary for me to be able to take notice of myself as others do, to listen to myself talking, to be able to draw conclusions from what I say! (PI, p. 192a)

In order to self-ascribe a belief on the basis of self-observation as one ascribes beliefs to others on the basis of observing them, one would have to observe oneself – listen to oneself speaking, for example – from a third-person perspective.

But this one cannot do, for one's beliefs are manifested in one's *intentional* actions (cf. LPP 67; PI, p.192b; RPP I §712). To take the stance of an observer towards one's own actions involves treating them either as accidental, and thus as not expressive of any belief, or else as expressive of the intentions, and thus the beliefs, of some other agent, as it were, 'inhabiting one's body'.

To say 'I believe that *p*' on the basis of self-observation and thus to use this phrase similarly to '*He* believes that *p*', as a description, implies a split personality (RPP I §820; Schroeder, S. 2006b, p. 171):

> If I listened to the words of my mouth, I might say that someone else was speaking out of my mouth.
>
> "Judging from what I say, *this* is what I believe." Now, it is possible to think out circumstances in which these words would make sense.

> And then it would also be possible for someone to say “It's raining and I don't believe it”, or “It seems to me that my ego believes this, but it isn't true.” One would have to fill out the picture with behaviour indicating that two people were speaking through my mouth. (PI, p. 192d-e)

One could at most infer what one *did* believe from one's *past* behaviour.[81] In order to infer one's *present* beliefs, one would have, as it were, to wait and see what one intentionally said and did when put in a situation which would 'test' whether one had the belief, e.g. on being asked 'Do you believe that *p*?', which is an absurd notion. In 'observing' one's own intentional actions, one would necessarily interfere with what one was trying to observe, as one would have intentionally to refrain from influencing one's own behaviour, i.e. from acting (or refraining from acting) intentionally. But it was precisely one's own intentional actions that one was trying to observe!

Furthermore, if one's present beliefs are phenomena which one discovers through self-observation and thus about which one can doubt or be in error, then doubts and error with regard to one's own present intentions would have, similarly, to be possible (cf. PI, p. 191g). For one's intentions *involve* beliefs[82] (Schroeder, S. 2006b, pp. 171-4). However, whilst one may doubt or be in error regarding whether or not one will succeed in carrying out one's intentions,[83] one can hardly doubt or be wrong about what one's present intentions *are* (although one may, of course, be undecided in one's intentions.)

Thus, by highlighting the conceptual connection between one's beliefs and one's intentional actions, Wittgenstein dispels the notion that 'I believe…' is a description of me as 'He believes…' is a description of another person.

81 However, even this much is only possible where one has forgotten what one then believed and has somehow been informed of one's behaviour at the time.

82 Even one's intention to stand up, for example, involves the beliefs that one *can* stand, that the floor will support one's weight, etc.

83 Or even (where there is time for one to change one's mind), whether one will *attempt* to carry out an intended action. (Schroeder, S. 2006b, p. 174)

As noted, in place of the misconceived notion that 'I believe that *p*' describes a state of mind, Wittgenstein suggests that it is a manifestation of one's belief, as is the simple assertion '*p*'. Indeed, Wittgenstein claims that

> …"I believe p" means roughly the same as "p"… (RPP I §472)

To say 'I believe that *p*' is very often a hesitant way of asserting '*p*' (Malcolm 1995, p. 197; Schroeder, S. 2006b, p. 162; cf. RPP §821; PI, p. 192l-m). In such a case, saying 'I believe that *p*, but not-*p*' is equivalent to saying

(19) Probably *p*, but not *p*.

Alternatively, 'I believe that *p*' may be an *emphatic* assertion that *p* (PI, p. 191b; Malcolm 1995, p. 197; Schroeder, S. 2006b, p. 162). In such a case, the utterance 'I believe that *p*, but not *p*' is equivalent to the utterance

(20) Definitely *p*, but not-*p*.

Again, 'I believe that *p*' may be a polite expression of disagreement. Thus the assertion 'I believe that *p*, but not *p*' is exposed as equivalent to

(21) With respect, *p*, but not-*p*.

In all of these cases, the seemingly paradoxical assertion is exposed as a disguised contradiction – a piece of patent nonsense – and thus the air of paradox dissolves.

It may, however, be objected that *of course* 'I believe that *p*' may be used to describe myself or to tell someone else about my belief and not to inform them of the fact that *p*. To suggest otherwise is simply to fly in the face of the obvious facts. And that it is *this* use of 'I believe that *p*' which is puzzling.

Aware of this objection, Wittgenstein responds:

> "I believe…" throws light on my state. Conclusions about my conduct can be drawn from this expression…
>
> If, however, "I believe it is so" throws light on my state, then so does the assertion "It is so". For the sign "I believe" can't do it, can at most hint at it. (PI, p.191c-d)

'I believe that *p*' *can*, of course, be said to give information about the speaker. However, the bare assertion '*p*' achieves this no less. In fact, to make the bare assertion is often the *more* appropriate way of communicating information to others about oneself. In an examination, for example, the pupil makes various assertions, but these serve the purpose of telling the examiner something about the pupil and not about the subject matter (PI, pp. 190m-191a).

The prefix 'I believe' *may* indeed serve to indicate that the *purpose* of one's assertion is to enlighten one's audience with regard to one's disposition. This, however, is a mere conversational expedient, which can be dispensed with in contexts, such as an examination, in which it is well understood that it is features of the speaker which are of interest, rather than the subject matter of his assertions.

Furthermore, the implicature that the purpose of 'I believe that *p*' is to inform one's audience about oneself can be cancelled without any absurdity, as can the implicature that the purpose of '*p*' is to inform one's audience about the subject-matter of the assertion. In neither case, however, can one retract the implication that one believes that *p*.

To the extent that 'I believe' may serve such a purpose, 'I believe that *p*' may be called a self-description, but the differences between this use of 'description' and its use in other cases must be borne in mind if philosophical confusion is not to ensue.

Sentence Meaning and Utterance Meaning: A Crucial Distinction

In claiming that 'I believe that p' means the same as '*p*', however, it may seem as though Wittgenstein propounds just the kind of non-moral subjectivism held up to ridicule by the Parity Thesis.

The grounds for Wittgenstein's claim, however, go back to his reminder that meaning is use (cf. PI §43), and we can distinguish between 'having a use' and 'being used'. In the *Investigations*, Wittgenstein invites us to

> Look at the sentence as an instrument, and at its sense as its employment.
> (PI §421)

An instrument or a tool can be put to a certain use on a particular occasion, but it also *has a use* independently of *being used*. A hammer may *be used* e.g. to drive a nail, on a particular occasion, but it also *has a use* (or uses) independently of its being so used, in that it *can be used* e.g. to drive a nail or to shape a piece of metal (and not, say, to turn a

screw).[84]

Similarly, a sentence has a use independently of its being used on a particular occasion in that it can be used to make various meaningful utterances,[85] including as part of an utterance of a larger, molecular sentence. However, a particular utterance of a sentence also has a meaning, or a sense, in virtue of its being used to do something on that occasion.

Thus, we can distinguish between the meaning of a sentence and the meaning of an utterance. Two sentences have the same meaning if they *can* be used in the same ways, whereas two utterances have the same meaning if they *are* used in the same way.[86]

Two sentences may be used in the same way on a particular occasion, and thus produce utterances which have the same meaning, without it being the case that they can be used in the same way in *every* context and thus have the same meaning as sentences. On the other hand, two utterances of a sentence with the same meaning can, when made in different contexts, have different meanings.[87]

The sentences 'I believe that *p*' and '*p*' are put to roughly the same standard assertive use i.e. to express the belief that *p*. When uttered as part of logically more complex sentences, however, the uses of these sentences diverge.

Thus, as utterances, 'I believe that *p*' and '*p*' have the same meaning (are used in the same way: to express the belief that *p*) but as sentences they do not have the same meaning, for they cannot be used interchangeably in *every* context.[88]

In his discussion of Moore's Paradox, Wittgenstein repeatedly comes back to a

84 This is, of course, only an analogy. That a hammer can't be used to turn a screw is a matter of empirical fact. Some clever person might, after all, come up with a way to use a hammer to turn a screw. No matter how clever one is, though, one cannot, as a matter of logic, use a word or sentence to do something which, by convention, it cannot be used to do (unless, of course, one stipulates a new sense for the expression, as when one uses a code, for example.)

85 NB - A speaker can invent and use indefinitely many type-sentences, but cannot similarly invent and use novel words without first establishing a convention for their use. Thus it is primarily words which have a meaning by convention (Baker & Hacker 1980a, p. 76).

86 This is not to deny that there may, indeed, be aspects of the use of an utterance which are not part of its literal meaning e.g. the use of a sentence to effect a conversational implicature. Thus, we might say that the literal meaning of an utterance is the way it is standardly or conventionally used in a particular context, although a speaker can convey something beyond what an utterance literally means.

87 This is, perhaps, most obviously exemplified by sentences containing indexical elements.

similar observation:

> Moore's paradox can be put like this: the expression “I believe that this is the case” is used like the assertion “This is the case”; and yet the *hypothesis* that I believe this is the case is not used like the hypothesis that this is the case.
>
> So it *looks* as if the assertion “I believe” were not the assertion of what is supposed in the hypothesis “I believe”!

> Similarly: the statement “I believe it's going to rain” has a meaning like, that is to say a use like, “It's going to rain”, but the meaning of “I believed then that it was going to rain”, is not like that of “It did rain then”. (PI, p. 190)

> Think of the expression “I say....”, for example in “I say it will rain today”, which simply comes to the same thing as the assertion “It will....”. “He says it will....” means approximately “He believes it will....”. “Suppose I say....” does *not* mean: Suppose it rains today. (PI, p. 192)

That tokens of sentences which differ in meaning can nonetheless have the same (utterance) meaning when asserted is what engenders Moore's Paradox.

88 The meanings of these sentences do, however, overlap in that they have more or less the same standard assertive use. Thus, although we might hesitate to say that they are *synonymous*, we might say that their meanings are the same in at least one important respect.

In fact, there is no reason to suspect that *any* two sentences can be used interchangeably in *all* contexts. However, the use of the sentences 'ϕing is (morally) wrong' and 'One ought (morally) not to ϕ', for example, overlap a great deal more than do 'ϕing is wrong' and 'I disapprove of ϕing'. Thus, we are inclined to call former, but not the latter pair 'synonymous' (although we may be inclined to call the former only 'near synonyms' were we to think of contexts in which they would not be intersubstitutable without a change of utterance meaning).

The point can also be put in terms of the verb 'to believe'. This verb is not put to the same use in simple first-person present utterances as it is, for example, in third-person or past tense utterances (cf. Schroeder, S. 2006b, p. 164). In first-person present utterances, this verb may have *no* real function (cf. Schroeder, S. 2006b, p. 166; PI, p. 192j),[89] whereas in its third-person use it serves the purpose of describing a person's disposition.

As Wittgenstein recognised, however, this will likely raise the objection:

> "But surely 'I believed' must tell of just the same thing in the past as 'I believe' in the present!"…(PI, p.190)

Wittgenstein, though, responds that

> Surely √-1 must mean just the same in relation to -1, as √1 means in relation to 1! This means nothing at all. (PI, p.190)

A word needn't serve the same function in every sentential context any more than a numeral need serve the same function when used in conjunction with different mathematical operators. The square root of minus one is an imaginary number, not a numeral that can be written down to so many decimal places, whereas the square root of 1 is 1. Indeed, a numeral performs no function at all when what is produced by using it in conjunction with a certain operator is nonsense (as when one divides by zero.)

'I believe that *p*', however, isn't nonsense, and it will be objected that the word 'believe' must have the same function in this simple expression as it does in, say, the conditional 'If I believe that *p*…', for arguments of the form

89 Other than, perhaps, as a pragmatic device for indicating that the *purpose* of one's expression of belief is to reveal one's disposition to one's audience.

(P1) If I believe that *p,* then *q*

(P2) I believe that *p*

(C) *q*

are clearly valid, whereas if (P2) did not assert what was hypothesised in (P1), then surely the argument would be invalid.

It is true that (P2) does not serve to describe my disposition, whereas what is hypothesised in (P1) is indeed my having a certain disposition. However, although these two tokens of 'believe' are put to different uses, they are still tokens of a type-word with a particular use, and this use is precisely what determines that the word will be used differently in these two contexts (cf. PI, p. 192h; RPP I §§490f). Thus there is no equivocation, or at least none of the kind that would invalidate the inference.

But, it may be asked, what accounts for the validity of this argument if what is hypothesised in the major premise is not asserted in the minor premise? The simple answer is that to express the belief that *p* by saying 'I believe that *p*' *is* to assert what is hypothesised in the conditional 'If I believe that *p*…', even though to express a belief through such a self-ascription is not to describe oneself as having it. Expressing the belief that *p* by saying 'I believe that *p*' *is* what we call 'asserting that one believes that *p*'.

There is good reason, though, why we consider such arguments valid, or why an inference from the premises to the conclusion of such an argument is taken as licensed. If you believe that, *if* I have a musket, you should hand over your gold, and I now produce a musket, then I have given you no less (and perhaps rather more) reason to believe that I have a musket, and that you should therefore hand over your gold, than had I described myself as having a musket.

Of course, my producing the musket does not represent a premise in a valid

argument, because arguments are essentially linguistic entities. The argument above, on the other hand, is formally valid, but it is not, for all that, dissimilar to the musket case. For someone who accepts the major premise (i.e. believes 'If *he* believes that *p*, then *q*'), a sincere expression of belief on my part is a perfectly good reason to accept '*q*'. Expressing my belief by way of the avowal 'I believe that *p*' is analogous to producing the musket.[90] In making this assertion, I manifest and thus reveal that I have the disposition that is hypothesised in the antecedent of the conditional major premise. In this case, however, I do this through the use of a sentence and the very same sentence, with the same conventional use and thus meaning, that appears in the antecedent of the conditional major premise, even if the use to which it is put on that occasion (as part of a different *utterance*) is different.

An Aside Concerning Meaning and Use

The distinction between sentence meaning and utterance meaning is a crucial one for my argument in this thesis. The idea, however, that Wittgenstein employed such a distinction, and even that such a distinction can be made, has been brought into question by, amongst others, some prominent Wittgensteinians. It is therefore worth pausing at this juncture to examine the distinction in finer detail, in order to rebut such claims.

The concept of linguistic meaning was central to Wittgenstein's thought throughout his career. The reason for this preoccupation is fairly simple. Philosophy traditionally aims at truth. But questions of sense are antecedent to questions of truth. For one can't even *begin* to go about settling the question of a proposition's *truth* if its *sense* is not yet clear. In

90 Indeed, you might even doubt that my expression of belief is genuine, as you might doubt that what I have produced is a really a musket.

order to settle questions of sense, though, one must know what *determines* a proposition's sense.

In the *Tractatus*, Wittgenstein made fateful assumptions about the nature of a proposition. He considered a proposition to be essentially a tool for representing reality, which it can do either correctly or incorrectly. A proposition achieves this feat in virtue of its being (ultimately, after proper philosophical analysis) composed of simple names and logical operators, combined in accordance with the rules of 'logical syntax', which mirror the possible ways in which the simple objects of which reality is ultimately composed can be combined.

Metaphysicians, though, regularly offend against the rules of logical syntax by, for example, employing 'formal concept-words' as 'proper concept-words' (e.g. in attempting to prove that 'there are objects', a statement which does not represent a particular configuration of objects), and ascribing 'internal properties' to things (e.g. in saying that 'the table is an object', which is not a contingent, bipolar proposition), and in doing so, according to the *Tractatus*, they talk nonsense (Baker & Hacker 1985, p. 41; TLP 3.323f, 4.003).

Wittgenstein, however, soon noticed “grave mistakes” (PI, p. viii) in his early account of language. He began to notice that language could be used to do much more than simply to depict states of affairs. In line with this recognition, Wittgenstein changed his mind about what meaning was, and in what meaningfulness consisted, but the insight that questions of sense are antecedent to questions of truth remained with him.

At the heart of Wittgenstein's later philosophy is the claim that

> For a *large* class of cases – though not for all – in which we employ the word "meaning" it can be defined thus: the meaning of a word is its use in the language. (PI §43)

The claim is, perhaps, one of Wittgenstein's most intuitively plausible and is disarmingly straightforward to establish. For, what one learns in learning the meaning of a word is precisely how to use the word correctly (Glock 1996, p. 206). It would be absurd to claim that someone knew the meaning of a word if he was unable to use it correctly, and equally absurd to claim that he did not understand it, even though his use of the word was impeccable.[91] And what goes for words goes, *mutatis mutandis*, for sentences and utterances. The meaning of a sentence or utterance is its use (PI §421; Baker & Hacker 1980a, p. 74).

The motto that 'meaning is use', however, is ripe to be misunderstood, and one particular misunderstanding of Wittgenstein's dictum, as applied to sentences, is particularly instructive for our purposes.

The misunderstanding in question is, unfortunately, fairly widespread. John Searle provides us with a particularly stark example:

> Wittgenstein's chief message about meaning was that meaning is use…One ambiguity in this approach was that it failed to distinguish between meaning as a matter of established conventional usage and the meaning of a

91 This is not to say that there are not those who have sought to cast doubt on Wittgenstein's equation of meaning and use (see e.g. Rundle 1990, Chs. 1, 9-10; Lycan 2000, pp. 94-7; Searle 1969, pp. 136-41; Grice 1989; Chs. 1-3; Grayling 1988, pp. 99f). Whilst some of these attacks certainly succeed in establishing that not *every* aspect of a word's use is relevant to its meaning, this point was explicitly conceded by Wittgenstein. It is only when Wittgenstein's dictum is taken as a philosophical theory, rather than as a trivial grammatical reminder, that such criticisms seem to have any force (Schroeder, S. 2006a, pp. 168-85; cf. Glock 1996). Such observations should not be seen as objections to a theory, but as further contributions to the grammar of the word 'meaning', that is, they tell us what aspects of use are and are not relevant to meaning (Baker & Hacker 1980b, pp. 118-21; cf. RPP I §633). This point is discussed in further detail in Chapter Five.

> particular utterance on a particular historical occasion. In short, it failed to distinguish between sentence and word meaning on the one hand, and speaker or utterance meaning on the other hand. (Searle 2007, p. 10)

Indeed, allegedly Wittgensteinian 'semantic contextualists/occasionalists' embrace the notion that no such distinction can be made. As Emma Borg (a 'minimalist'[92] opponent of 'occasionalism') writes:

> …minimalism holds that there is a meaning for sentences…Occasionalism, on the other hand, holds that meaning emerges only within a context of use…(Borg 2010, p. 103)

The claim can also be found in the writings of other Wittgensteinians. Thus, David Macarthur has claimed that

> …the term "use" in Wittgenstein refers to something quite specific, someone or other's use of words on a certain occasion for a certain purpose. (Macarthur 2010, p. 93)[93]

92 Minimalism, however, makes the opposite mistake, in denying, or at least downplaying the significance of utterance meaning.

93 The notion that meaning is a feature only of particular linguistic acts also underlies the so-called 'New Wittgenstein' movement, a central claim of which is that there can be no such thing as nonsense that is made up of meaningful words illegitimately combined, since it is only in the context of a meaningful sentence that a word has a meaning (see e.g. Diamond 1981). This reading of Wittgenstein has been roundly criticized precisely on the grounds that it fails to distinguish between the conventional use of a word, and the use to which it is put on a particular occasion (Glock 2004). Whilst this is an interesting topic in its own right, the matter requires much more attention that I can give it here. Moreover, The 'New Wittgensteinian' reading is concerned more with the distinction between the meaning of a token use of a word and that of the conventional use of a word (New Wittgensteinians being committed, absurdly, to the notion that there is no such thing as the latter, or at any rate, that a token use of a word in a nonsensical utterance does not have a meaning.) Whilst these concerns are interestingly related to the debate regarding sentence and utterance meaning, they are not directly relevant to my argument.

Does Wittgenstein Deny Sentence Meaning?

As Gordon Baker & Peter Hacker point out

> …Wittgenstein does not explicitly draw the distinction between the meaning of a sentence and what is said by the use of a sentence on a given occasion. He makes the word 'Satz' do far more work than a single term can comfortably do, although one must admit that what he has in mind is almost always perfectly clear. (Baker & Hacker 1980a, p. 73)

It is true that Wittgenstein was often not careful to distinguish between what I call 'sentence meaning' and 'utterance meaning' and sometimes gave the impression that there was only the latter. It is not surprising that Wittgenstein did not keep a great distance between the concepts 'sentence', 'utterance', 'proposition' and the like. After all, what one utters in making an utterance, or in expressing a proposition typically *is* a sentence. It was, nonetheless, remiss of Wittgenstein to be so lax in his terminology. For, as Searle rightly points out elsewhere (Searle 1978, p. 209), it is a category mistake to confuse even a token utterance of a sentence with 'what is said' by the use of a token sentence.

Perhaps the most likely candidates for passages supporting the notion that Wittgenstein held that meaning is a feature only of utterances in a particular context are to be found in *On Certainty*:

> I know that a sick man is lying here? Nonsense! I am sitting at his bedside, I am looking attentively into his face. – So I don't know, then, that there is a sick man lying here? Neither the question nor the assertion makes sense.

> Any more than the assertion “I am here”, which I might yet use at any moment, if suitable occasion presented itself. Then is “2 x 2 = 4” nonsense in the same way, and not a proposition of arithmetic, apart from particular occasions? “2 x 2 = 4” is a true proposition of arithmetic – not “on particular occasions” nor “always” – but the spoken or written sentence “2 x 2 = 4” in Chinese might have a different meaning or be out and out nonsense, and from this is seen that it is only in use that the proposition has its sense. And “I know that there's a sick man lying here”, used in an *unsuitable* situation, seems not to be nonsense but rather seems matter-of-course, only because one can fairly easily imagine a situation to fit it, and one thinks that the words “I know that...” are always in place where there is no doubt, and hence even where the expression of doubt would be unintelligible. (OC §10)

> “I know that that's a tree.” Why does it strike me as if I did not understand the sentence? Though it is after all an extremely simple sentence of the most ordinary kind? It is as if I could not focus my mind on any meaning. Simply because I don't look for the focus where the meaning is. As soon as I think of an everyday use of the sentence instead of a philosophical one, its meaning becomes clear and ordinary. (OC §347)

It ought to be noted, however, that the contrast set up here is not between sentences in use, and sentences as such, but between suitable and unsuitable uses of sentences. Whilst it is true, then, that Wittgenstein's concern in these passages is with utterance meaning, there is nothing in these passages that serves to deny that sentences also have a meaning.

Indeed, this should not be surprising, for many passages seem flatly to contradict the claim that Wittgenstein overlooked the distinction between sentence meaning and utterance meaning:

> "Grasshoppers don't think." Where does this belong? – Is it an article of faith, or does it belong to natural history? If the latter, it ought to be a sentence something like: "Grasshoppers can't read and write." This sentence has a clear meaning, and even though it is perhaps never used, still it is easy to imagine a use for it. (RPP II §24)

> The fact that we understand a sentence shows us that we could use it in certain circumstances (even if it were only in a fairy tale), but this does not show us *what* we can do with it and *how much*. (LW I §383)

> But the meaning of the sentence "There are humans who see", i.e. its possible use, is not immediately clear at any rate. (LW II, p. 77)

Wittgenstein, then, was not unaware of the distinction between sentence and utterance meaning. In fact, he had explicitly talked about such a distinction as early as 1930, (Moore 1954, pp. 10f) and the distinction was also implicit in the *Tractatus* notion of 'sign vs. symbol' (cf. Black 1964, pp.117f).

Is Sentence Meaning Potential Utterance Meaning?

A natural way to think of sentence meaning is as potential utterance meaning. As well as supporting the idea that not only particular utterances, but also sentences apart from a context of utterance have meaning, the above passages would appear to lend credence to such an account of the meaning of a sentence.

Ironically, given Wittgenstein's apparent endorsement of it, this is the view of sentence meaning that Searle himself recommends:

> …the best way to understand the meaning of a sentence is to see it as a potential utterance. We understand the meaning of the *sentence* "It's raining" because we understand that it can be, in appropriate circumstances, used to make the *statement* that it is raining. The sentence is to its use as the tool is to its use. You can have a hammer that you don't hammer with, but you don't understand what a hammer is if you don't know what hammering is. Analogously you can have a sentence that you never use to say anything, but to understand the sentence is to understand what it could be used to say. (Searle 2007, p. 10)

A number of other passages, too, support the interpretation that, far from overlooking the distinction between sentence and utterance meaning as Searle claims, Wittgenstein in fact endorsed Searle's recommendation to conceive of sentence meaning in terms of potential utterance meaning. In some passages, for example, Wittgenstein appears to suggest that a sentence has a meaning in virtue of our being able to think of a suitable use for it in making an utterance, and that a sentence can fail to make sense because we

cannot think of a suitable situation in which it might be used to make a meaningful utterance:

> "After he had said this, he left her as he did the day before." – Do I understand this sentence? Do I understand it just as I should if I heard it in the course of a narrative? If it were set down in isolation I should say, I don't know what it's about. But all the same I should know how this sentence might perhaps be used; I could myself invent a context for it.
>
> (A multitude of familiar paths lead off from these words in every direction.) (PI §525)

> If someone were to say "Red is composite" – we should not be able to guess what he was alluding to, what he will be trying to do with this sentence. But if he says: "This chair is composite" although we may not know what kind of composition he is speaking of, still we can at once think of more than one sense for his assertion.
>
> Now what kind of fact is this, that I am drawing attention to here?
>
> At any rate it is an *important* fact. – We are not familiar with any technique, to which that sentence might be alluding. (RPP I §605)

As our discussion of Moore's paradox shows, however, the possibility of using a sentence on its own as an utterance is not a *necessary* condition of its meaningfulness. For there can be perfectly meaningful sentences which are plain nonsense when asserted, but which can nonetheless be embedded in more logically complex sentences to form perfectly intelligible utterances. An example of such a sentence is (cf. PI, p. 190; LW I §141):

(22) I believe falsely that *p*

It makes no sense to say (22), but it makes perfect sense to say:

(23) If I believe falsely that *p*, then *q*

or:

(24) It's possible that I believe falsely that *p*

Hence (22) has no potential use in making an utterance when used atomically, but nonetheless has a potential use in embedded contexts.[94]

Moreover, whilst it may be a *sufficient* condition of a sentence's being meaningful that it has a potential use in making an utterance, understanding this potential use is *not* sufficient for understanding the sentence. Thus, Wittgenstein suggests that on the lips of a child who has learned to say 'I have a pain', but cannot yet say 'When I have a pain, the Doctor comes', the former sentence (and, indeed, its constituent words) have a different meaning than they do in the mouth of a mature English speaker (LW I §899). Thus, the child may understand perfectly well that this string of words can be used to express his pain, he cannot be said yet to understand the sentence as an adult does. At best he understands a sentence in his own idiolect, in which it has only this simple assertive use. Thus, once again, sentence meaning often outstrips the potential utterance meaning of that sentence on its own.

94 In fact, one can think of a use for nearly any utterance if one tries hard enough. Wittgenstein thought of a use for (22), imagining a situation in which the person who utters it has a kind of 'split personality'. However, we are able to understand (22) perfectly well – to use it, for example, in the antecedent of a conditional – even if we haven't the imaginative powers to think up such a use, thus its meaningfulness is not based on our ability to find a simple assertive use for it.

Sentence meaning, then, is not simply potential utterance meaning. Rather, the meaning of a sentence, like that of a word, is its whole use, or what it *can* be used to do. In addition to its simple assertive use, it has a use in contributing to logically more complex sentences, and this too is part of its meaning, carried with it into the various contexts in which it is employed, much as a claw hammer, whilst being used to hammer in a nail, is still a tool which *can* also be used to pull a nail out again. In many passages, Wittgenstein is quite explicit on this point

> One may have the feeling that in the sentence "I expect he is coming" one is using the words "he is coming" in a different sense from the one they have in the assertion "He is coming". But if it were so how could I say that my expectation had been fulfilled? If I wanted to explain the words "he" and "is coming", say by means of ostensive definitions, the same definitions of these words would go for both sentences.
>
> But it might now be asked: what's it like for him to come? – The door opens, someone walks in, and so on. – What's it like for me to expect him to come? – I walk up and down the room, look at the clock now and then, and so on. – But the one set of events has not the smallest similarity to the other! So how can one use the same words in describing them? – But perhaps I say as I walk up and down: "I expect he'll come in" – Now there is a similarity somewhere. But of what kind?!
>
> It is in language that an expectation and its fulfilment make contact. (PI §§444f; cf. PG, p. 59)

Here, Wittgenstein says that the words 'he is coming' have the same meaning in the

assertion 'He is coming' and in the expression of expectation 'I expect he is coming'. Yet the uses to which each of these occurrences of the sentence is put are quite different.

Likewise, on several occasions, Wittgenstein gives examples of cases in which two sentences are put to the same use, but, in general, have different meanings:

> The question whether “He can continue...” means the same as “He knows the formula” can be answered in several different ways: We can say “They don't mean the same, i.e., they are not in general used as synonyms as, e.g., the phrases 'I am well' and 'I am in good health'”; or we may say “*Under certain circumstances* 'He can continue...' means he knows the formula”. (BB, p. 115)

And similarly

> But did “Now I can go on”…mean the same as “Now the formula has occurred to me” or something different? We may say that, in those circumstances, the two sentences have the same sense, achieve the same thing. But also that *in general* these two sentences do not have the same sense. (PI §183)

A Potential Problem?

It may, however, be objected that if a sentence can be meaningful simply in virtue of its being useable in an embedded context, then *any* syntactically declarative sentence is meaningful. After all, it is an important part of Wittgenstein's method that it's possible to

'operate' with nonsense, in order to transform it from 'latent nonsense' into 'patent nonsense' (Glock 2004, pp. 237f; cf. PI §§464, 564). It is essential to this kind of operation that nonsensical sentences can be put in both the antecedent and the consequent of a conditional e.g. 'If the meaning of a word is its bearer, the meaning of 'Mr. N.N.' can die' (PI §40). Thus, even a piece of nonsense, so long as it is in the form of a syntactically declarative sentence, can be put to use in a *reductio ad absurdum* argument. Indeed, it might be argued, *any* piece of nonsense '*n*' can be given a use simply by embedding it in a sentence of the form 'It is nonsensical to say *n*'.

Firstly, one might counter that, in calling a sentence 'nonsensical', one mentions, and does not use the sentence. Moreover, to say that a word or phrase has a use on account of the fact that it can be used in order to point out its own uselessness is like saying 'Even a broken tool has a use, i.e. as an example of a broken tool'. It would not be arbitrary to restrict the use that constitutes meaning to uses other than in pointing out that the sentence is useless!

Wittgenstein, then, did distinguish between sentence meaning and utterance meaning. Indeed, such a distinction is, in fact, suggested by his dictum that meaning is use, rather than overlooked by it. A sentence is meaningful, not simply in virtue of having the potential to be used atomically in order to make a meaningful utterance, since there are meaningful sentences which have no such simple use, but also in virtue of its possible use in more logically complex contexts.

The Parity Thesis Revisited

From the foregoing, we might hazard an explanation of the paradoxical nature of

statements of the form

(1) *x* is wrong, but I don't disapprove of *x*.

or

(2) *x* is not wrong, but I disapprove of *x*.

Wittgenstein's arguments regarding avowals of belief apply *mutatis mutandis*, to avowals of attitude. Although it is tempting to say that 'I disapprove of *x*' describes a speaker or his attitudes, this claim does not stand up to scrutiny. Were it so, (1) and (2) would not be paradoxical.[95]

Attitudes of approval and disapproval are, like beliefs, dispositions of a special kind, in that a speaker's avowal that he has such an attitude is a manifestation of that attitude and not a fallible self-description. A speaker cannot take a third-person stance towards his own attitudes of approval and disapproval, ascribing them to himself on the basis of self-observation, any more than he can take such a distanced attitude towards his own beliefs. To do so would be to treat his actions and avowals, and therefore the attitudes which they betray, as those of a third-party.

Moore's Paradox is overcome in the moral case, just as in the non-moral case, only if we instead recognise that the function of 'I disapprove of *x*' is to *express* disapproval towards *x*.

Thus, on an expressivist account, the meanings of the *utterances* 'I disapprove of *x*' and '*x* is wrong' are explained in the same way: they are both expressions of disapproval towards *x*. This does not, however, commit the expressivist to the biconditional '*x* is wrong

95 As the expressivist claims they are.

iff I disapprove of *x*'. This biconditional makes the claim that '*x* is wrong' and 'I disapprove of *x*' can also be used interchangeably when embedded in more logically complex sentences, i.e. that they have the same meaning as *sentences*. But, as we saw in the non-moral case, that two sentences have the same standard assertive use does not entail that they are synonymous as sentences.

The Frege-Geach Problem

All of this, however, leads the expressivist directly into the jaws of the infamous Frege-Geach problem:

> A thought may have the same content whether you assent to its truth or not; a sentence may occur in discourse now asserted, now unasserted, and yet be recognizably the same sentence. (Geach 1965, p . 449.)

Peter Geach takes the above observation, which he calls the 'Frege Point' (Geach 1965, p. 449),[96] to lead to a decisive refutation of non-cognitivism. In order to bring out the allegedly fatal consequences for non-cognitivism of the Frege Point, Geach presents the following argument:

> If doing a thing is bad, getting your little brother to do it is bad.
>
> Tormenting the cat is bad.
>
> *Ergo*, getting your little brother to torment the cat is bad.

96 After Gottlob Frege, to whom he credits the original insight.

Geach goes on, quite rightly, to point out that

> The whole nerve of the reasoning is that “bad” should mean exactly the same at all four occurrences – should not, for example, shift from an evaluative to a descriptive or conventional or inverted-commas use. But in the major premise the speaker (a father, let us suppose) is certainly not uttering acts of condemnation: one could hardly take him to be condemning just *doing a thing*. (Geach 1965, pp. 463f.)

The upshot of Geach's observations is that, as the thesis that moral terms serve to express speakers' attitudes, non-cognitivism seems embroiled in a dilemma. Either the non-cognitivist must deny the 'Frege Point' and insist that *every* use of a particular moral term serves to express the very same attitude, or else he must accept that, in unasserted contexts, moral terms undergo a change of meaning, rendering fallaciously equivocal apparently valid arguments containing such terms. Either way, non-cognitivism is reduced to absurdity.

Geach's dilemma, however, is a false one: non-cognitivists are committed to neither of these absurd claims. To take the non-cognitivist as committed to denying the Frege Point would be uncharitably to misconstrue his position as stronger than it is. The non-cognitivist's claim, in Geach's own words, that “To call a kind of act bad is not to characterize or describe that kind of act but to condemn it” (Geach 1965, p. 462) does not commit him to the further, obviously false, claim that *every* use of the term 'bad' constitutes a speech-act of condemnation or expression of disapproval (cf. Searle 1969, p. 138). After all, to say '*If* this is bad, then…' is *not* to *call* anything 'bad'!

It does not follow, however, that the non-cognitivist is committed to the

consequence that moral terms change their meanings in unasserted contexts. In arguing to this effect, Geach assumes that a difference of meaning (or of speech-act performed) between two utterances entails a difference in the meanings of the words of which the utterances consist. This, however, is an obvious non-sequitur. 'You are six feet tall!' and 'Is it true that you are six feet tall?' are quite different utterances; different speech-acts are performed in each case. The former is used to make an assertion, or to express a belief, whereas the latter is used to ask a question. Nevertheless, both utterances contain the sentence 'You are six feet tall' and it would be absurd to suggest that the predicate 'six feet tall' therefore means something different in each case.[97] A difference in word-meaning simply does not follow from a difference in utterance meaning or speech-act performed and Geach provides no argument to the contrary.

It might, though, be objected that, since a word is put to a different use in a different utterance, it does have a different meaning there. This, however, is to overlook the fact that the word has the same conventional meaning in each case, and it is this – its potential use – which determines that it may be used in these different ways in different parts of a perfectly valid argument. In order to make the charge of equivocation stick, Geach needs to show that there is a difference in conventional meaning between occurrences of a word, not merely that they are put to different uses in different contexts.

More charitable, then, perhaps, than Geach's apparent[98] claim to have *refuted* non-cognitivism by showing that it entails something false, is the claim that non-cognitivism is *incomplete*, since it tells us only what is done with simple assertive uses of moral predicates. Indeed, the objection can be made stronger, in that the non-cognitivist owes a story about the uses of moral terms in embedded contexts which coheres with his story

97 Or else, as John Searle quite rightly points out, the assertion would not be an answer to the question (Searle 1969, p. 137).

98 There is some debate as to whether Geach took his point to be the implausibly strong one that I am imputing to him, or the weaker claim that I think the non-cognitivist needs to answer (Dreier 1996, pp. 31ff).

about the uses of such terms in their alleged simple assertive use.

Whilst expressivists may have explained the meaning of the utterance '*x* is wrong', this is insufficient to have given a *complete* account of the meaning of the word 'wrong' or of the sentence '*x* is wrong'. In order for them to have done so, an account of the meanings of other utterances in which this word or sentence might occur would have to follow from their account of the meaning of a moral utterance.

From the foregoing, however, we can see that no such account *can* follow. Contrary to appearances, 'I disapprove of *x*' is *not* a description of a speaker's state of mind, but an expression of disapproval. Thus, on an expressivist account, the utterance-meaning of '*x* is wrong' is explained in the same way as that of 'I disapprove of *x*', i.e. both are expressions of disapproval. Thus it is clear that the expressivist's story about the utterance-meaning of '*x* is wrong', does not suffice to give the meaning of the *word* 'wrong'. The assertive use of the word 'wrong' in expressing disapproval is only a part of its meaning, given that, in unasserted contexts, it does not serve the same function (cf. Geach 1965, pp. 463f). Nor does it suffice to give the meaning of the *sentence* '*x* is wrong'. The sentences '*x* is wrong' and 'I disapprove of *x*' cannot, after all, be used interchangeably in all contexts to make utterances with the same meaning ('If *x* is wrong, then…', does not mean the same as 'If I disapprove of *x*, then…', for example) (cf. Searle 1970, p. 138).

Of course, the expressivist does say *something* more about the use of the sentence '*x* is wrong'. He says that it stands to 'I disapprove of *x*' as '*p*' stands to 'I believe that *p*'. But nothing in his story shows why this is so, or even guarantees that this relationship *can* hold between these sentences. Indeed, from the foregoing, we might conclude that all that the 'Parity Thesis' really tells us is that, whilst '*x* is wrong' and 'I disapprove of *x′* have the same standard assertive use, their uses diverge in embedded contexts.

This does not show the expressivist's claim about moral utterance-meaning to be

incorrect. Nor, indeed, does it show it to be useless, even with regard to explaining the meaning of a moral word or sentence. After all, expressivism, if correct, does describe an *aspect* of the use of the sentence '*x* is wrong' (i.e. that when uttered literally as an atomic sentence, it serves to express disapproval of *x*) and of the use of the word 'wrong' (i.e. that when used literally in a case of simple predication, it serves to express disapproval) (cf. Glock 1996).

The real problem posed by the Frege-point is that, whilst non-cognitivism's explanation of the meanings of simple moral utterances may be quite correct, it opens up questions as vexed as those which it was meant to put to rest, about, for example, what use moral sentences have in embedded contexts, and how this is consistent with the non-cognitivist's story about their simple atomic use. As a tool for dispelling philosophical confusion, then, the non-cognitivist explanation of moral utterance meaning seems to break as much as it mends.

V

Giving Subjectivism Careful Consideration

As I argued in Chapter Two, many of Wittgenstein's early concerns bear a striking resemblance to later non-cognitivist qualms about subjectivism. Indeed, his earlier writings, once stripped of their attendant metaphysics and theory of language have a great deal in common with non-cognitivism, and whilst, in his later writings on ethics, Wittgenstein's emphasis moved away from traditional metaethical debates, he nonetheless appears to have taken for granted a broadly expressivist position.

Wittgenstein's thoughts in this regard, however, were embryonic at best, and contemporary expressivists, who have attempted to see the non-cognitivist project through to a greater extent than did Wittgenstein, have encountered major difficulties, not least in the form of the Frege-Geach problem, a solution to which, I have argued, cannot follow from their explanations of the meanings of simple moral utterances. Hence the non-cognitivist is left needing to supplement his story with an explanation of the uses of molecular utterances containing atomic moral sentences.

In this chapter, then, I begin by discussing a number of recent attempts to take on this challenge. My intent is not to provide a decisive refutation of the non-cognitivist's position, but merely to indicate how tortured his path has become, even to the extent that we may question whether, in following him, we will ever reach our philosophical goal of clarifying moral concepts.

Perhaps, though, the debate has rushed ahead too quickly, for, as I shall argue, it isn't at all clear that there is any particular *need* to resort to non-cognitivism, with its

attendant perplexities, in order to provide a satisfying account of the use of moral terms.[99] When doing philosophy, we must keep in mind our aim of dispelling philosophical puzzlement about the use of words (and eschewing the practical consequences of such puzzlements), rather than that of building a 'theory'. If we can get what we want out of a metaethic without resorting to the convolutions of non-cognitivism, then there is every reason to do so.

To this end, I propose to take a step back and look at the motivations for espousing non-cognitivism in the first place. I argue that these are not, perhaps, as compelling as they have often been taken to be, and that a return to a more simple-minded subjectivism (although not 'simple' subjectivism) is the preferred choice.

The Minimalist Manoeuvre

As I have argued, non-cognitivists owe an account of the use of moral sentences in embedded contexts which is consistent with their explanations of the meanings of simple moral utterances. Whilst most non-cognitivists accept this demand and set about meeting it, so-called 'minimalists' have attempted instead to circumvent the demand altogether. Whilst it may be true, according to the minimalist, that non-cognitivists have traditionally told us a story only about the simple assertive use of moral predicates, all that is needed in order to flesh out this story to the whole use of a predicate thus described, is a further claim to the effect that the predicate in question functions as a 'logical' one, and therefore embeds happily in various unasserted contexts (see e.g. Horwich 1993, p. 75).

99 A similar point was made by Michael Smith (1986) in relation to A.J. Ayer's emotivism. Smith saw Ayer's emotivism as a too-quick response to apparent problems with alternative accounts – a response that came with many of its own problems in turn. Whilst I don't ultimately agree with the direction Smith went, the view I eventually recommend has some formal similarities to his.

The fault in this quick way with the embedding problem has been exposed in quite ingenious fashion by James Dreier. Dreier (1996, pp. 42ff) asks us to imagine a fictitious predicate 'hiyo'. The simple assertive use of this predicate is explained as follows: to call someone 'hiyo' is to greet them. If the minimalist is right, then, simply by stipulating that 'hiyo' is a 'logical predicate', we should now know all there is to know about how this predicate embeds in logically more complex sentences. But, of course, we don't! Even with the minimalist's stipulation in hand, we have no idea what sentences such as 'If Bob is hiyo, then the stock market has fallen ten points' mean. Indeed, it makes no sense whatsoever to embed the predicate 'hiyo' in the antecedent of a conditional, even with an explanation of its use in contexts of simple assertion in hand.

To stipulate that a predicate which has been defined in terms of its simple assertive use is a 'logical' one is simply to insist that the predicate embeds in the same way as do predicates whose use we do not find problematic. But this assurance is hollow, for many *do* find it puzzling how a predicate whose meaning is characterised as the non-cognitivist characterises moral predicates functions in embedded contexts. Indeed, nothing in the non-cognitivist's story guarantees that such a predicate must embed any more happily than does the fictitious predicate 'hiyo'.

The 'Involvement' Account

Most contemporary non-cognitivists, therefore, take the problem more seriously than do the 'minimalists' (see e.g. Gibbard 2003, pp. 63ff). Rather than simply insisting that the predicates which they explain in terms of their simple assertive use do embed comfortably in unasserted contexts, they attempt to give a systematic account of what uses sentences

containing moral predicates actually have in the various unasserted contexts in which they can appear, which is consistent with their story about the simple atomic use of such predicates.

The first significant attempt by a contemporary[100] non-cognitivist to tackle this problem was Simon Blackburn's 'involvement' account. Blackburn was concerned to give an explanation specifically of the meanings of conditional sentences, in order to tackle Geach's concern that non-cognitivists had no account of the validity of *modus ponens*.

To this end, Blackburn argued that conditional moral sentences make the claim that the consequent is 'involved' in the antecedent (Blackburn 1993, pp. 126-9). Thus, Geach's argument can be reconstructed as follows:

> [Tormenting the cat is bad] involves [getting your little brother to torment the cat is bad]
>
> Tormenting the cat is bad
>
> *Ergo* Getting your little brother to torment the cat is bad

The problems with the 'involvement' account are many and various. One fairly glaring problem is that the account stands no chance of being generalized. Blackburn was responding specifically to Geach's worry regarding *modus ponens* and the non-congnitivist's inability to account for the meanings of conditional sentences. Other argument forms are entirely overlooked by this account, and, indeed, it would appear that the only way that other argument forms could be made to fit the mould of the involvement account would be to make them *all* into forms of *modus ponens*. The inference from 'P∧Q'

100 The first significant attempt before Blackburn's was arguably R.M. Hare's in *The Language of Morals* (Hare 1952, ch. 2). This approach essentially involved the notion that an imperative contained a descriptive 'sentence-radical', which accounted for its inconsistency relations, an idea which Wittgenstein criticized in Frege's work (PI §22; cf. Baker & Hacker 1980a, p. 67; 1980b, pp. 78-86).

to 'P', for example, would have to be licensed via the addition of a further premise '[P∧Q] involves [P]', a move which seems highly implausible (Schroeder, M. 2010a, p. 114).

Moreover, this simply doesn't appear to be a plausible analysis of the function of conditional statements in general, but only of *analytic* conditional statements (Schroeder, M. 2010a, pp. 112f). Indeed, 'involves', as Blackburn uses it, appears simply to *mean* 'entails'. But I may, for example, believe that 'If my parents believed that killing is wrong, then killing is wrong', whereas it cannot plausibly be claimed that 'My parents believed that killing is wrong' *involves* 'Killing is wrong' – that my parents' moral belief entails its own truth!

The 'Higher Order Attitude' Account

Blackburn's next attempt is known as the 'Higher Order Attitude' account. According to this account, molecular moral utterances serve to express 'higher order' attitudes of approval and disapproval[101] *towards* certain other attitudes or combinations of attitudes (Blackburn 1984, p. 195).

This can, once again, be explained by applying it to Geach's worrisome *modus ponens*. Blackburn accepts the standard non-cognitivist account of an atomic moral sentence, thus the minor premise

P1 Tormenting the cat is bad

101 In fact, Blackburn isn't entirely clear about what attitudes are expressed by moral sentences. (Unwin 1999, p. 339) He instead says that to say, e.g. 'It ought not to be the case that *p*' is to 'Boo!' *p* (in Blackburn's formalisation: B!*p*). As Blackburn's 'H!' and 'B!' operators are merely place-holders for as-yet-undetermined pro and contra attitudes, I take it that to substitute an actual attitude, such as disapproval is an acceptable way to clarify Blackburn's position. If the reader prefers, he can simply substitute Blackburn's operators for mention of any specific attitude.

expresses 'disapproval of tormenting the cat'. According to Blackburn, however, the conditional major premise,

P2 If tormenting the cat is bad, then getting your little brother to torment the cat is bad

expresses 'disapproval of [disapproval of tormenting the cat, but not of getting one's little brother to torment the cat]'.

Someone who has these two attitudes, it is argued, is under some rational pressure to go on to form the attitude expressed by the conclusion

C Getting your little brother to torment the cat is bad

namely 'disapproval of getting one's little brother to torment the cat'.

It is sometimes claimed that Blackburn's account fails to capture the seriousness of the irrationality involved in accepting the premises of a moral argument whilst denying, or failing to accept its conclusion. On Blackburn's account, it is argued, the irrationality of doing so consists simply in doing something that you yourself disapprove of, i.e. holding attitudes or combinations of attitudes that you yourself disapprove of, which, whilst it might make you in some sense conflicted, does not make you incoherent (Schroeder, M. 2010a, p. 118). Put differently, it is always possible for one to *do* what one disapproves of, and even to say so without thereby manifesting linguistic incompetence. But accepting the premises of an (obviously) valid argument, yet denying its conclusion would ordinarily constitute a criterion of having *misunderstood* some element of the argument (cf. Hare 1952, p. 25).

This objection, however, misses its mark. For it isn't at all obvious that it *is* possible to have an attitude of approval or disapproval of which you yourself approve or disapprove, as one can perform an action of which one disapproves, or fail to perform one of which one approves. One can, of course, disapprove of one's own tastes or preferences. I might, after all, disapprove of my own predilection for trashy American blockbusters, thinking that I ought, really, to try to make myself prefer high drama. But it isn't clear what it would *be* to take an attitude of disapproval towards one's own attitude of disapproval, if not simply to take a contrary attitude. To express disapproval of someone *else's* attitude would, after all, be taken as a manifestation of a contrary attitude, and there would appear to be no reason why this should not hold in one's own case too. Hence there is a very real incoherence in the notion that someone might disapprove of his *own* attitudes of approval or disapproval.

A more worrying criticism, then, is that such an account fails accurately to represent the inconsistency and entailment relations between moral utterances. For the following argument can be constructed, which, on Blackburn's account, should be logically equivalent to Geach's original *modus ponens*:

P1' Tormenting the cat is bad

P2' It's bad to disapprove of tormenting the cat and not to disapprove of getting your little brother to torment the cat

C' Getting your little brother to torment the cat is bad

On the Higher Order Attitude account P2' appears logically equivalent to P2 in the original argument: it expresses precisely the same attitude. Yet this argument seems invalid, whereas the original argument is clearly valid (Schroeder, M. 2010a, pp. 120f). Blackburn

might respond that, whilst the argument may not be valid by the rather crude standards of propositional logic, it would, nonetheless, be incoherent to accept the premises, but to deny the conclusion and that it is this fact that renders an argument valid, rather than the fact that one can get its tableau to close. To deny that this fails to capture the specifically 'logical' nature of the incoherence involved simply question-beggingly assumes an account of 'logic' that is already philosophically loaded in favour of the cognitivist. Our 'pre-philosophical' notion of this kind of incoherence is simply that there is something amiss with a person who reasons thus, and therefore it is an open question, ripe for philosophical inquiry, whether what is amiss is that we have expressed contradictory beliefs, or rendered ourselves unintelligible in some other way, such as by expressing mutually exclusive attitudes (Sinclair 2009, p. 143).

Nicholas Unwin, however, has made the stronger point that Blackburn's account[102] *can't* capture a fundamental part of the logic of our language, namely negation. According to Unwin (1999, p. 341), there is a “fundamental syntactic defect” in Blackburn's analysis of moral terms, in that there appears to be insufficient structure in this account of the meanings of moral sentences to make sense of the distinction between failing to accept a proposition and accepting the negation of a proposition.

In constructing his logic, Blackburn relies on the notion that the logic of belief mirrors the logic of the *contents* of belief. Beliefs are inconsistent, after all, if and only if their contents are inconsistent. The trouble, however, is that the non-cognitivist holds that to believe that *x* is wrong is to disapprove of *x*. Hence he accepts the equivalence 'I believe that *x* is wrong iff I disapprove of *x*.' But it turns out not to be possible to hold this equivalence, whilst maintaining that the logic of moral belief is explained by the logic of the *contents* of moral belief. For, take the following sets of sentences:

102 Unwin (2001) also makes a similar criticism of Allan Gibbard's attempts to tackle the embedding problem.

w Jon thinks that murdering is wrong.

n1 Jon does not think that murdering is wrong.

n2 Jon thinks that murdering is not wrong.

n3 Jon thinks that not murdering is wrong.

w* Jon disapproves of murdering.

n1* Jon does not disapprove of murdering.

n2* ???

n3* Jon disapproves of not murdering. (Schroeder, M. 2010a, pp. 135-6; Schroeder, M. 2008, p. 578.)

There are simply too few places in which to insert a negation into w*, compared with w, and, hence, it appears that there is no n2* equivalent of n2 to describe acceptance of a negation in terms of noncognitive attitudes, or at least none that allows Blackburn's semantics to be 'compositional', as he aims for it to be. After all, we *could* stipulate that n2* expresses an attitude towards murdering that is inconsistent with disapproval of murdering, for example, *approval* of murdering, or *indifference* towards murdering. But such a result does not fall neatly out of Blackburn's account, hence the negation problem at least gives Blackburn some explaining to do.

Bifurcated Attitude Semantics

In response to this last worry, Mark Schroeder (2008, p. 58) suggests that a moral utterance such as 'Tormenting the cat is bad' expresses the attitude of

(P1") Being for blaming for torturing the cat

In adding this extra level of structure to the attitude supposedly expressed in passing moral judgement, Schroeder appears to have overcome the 'negation problem' by providing an account of what is expressed by n2* above:

w' Jon is for blaming for murdering.

n1' Jon is not for blaming for murdering.

n2' Jon is for not blaming for murdering

n3' Jon is for blaming for not murdering.

Moreover, his account seems to capture the correct entailment and consistency relations. For example, the conditional premise 'If tormenting the cat is bad, then getting one's little brother to do so is bad', Schroeder (2008, p. 66) claims, expresses

(P2") Being for not [blaming for tormenting the cat and not blaming for getting one's little brother to do so]

One can't, according to Schroeder, be 'for' blaming for torturing the cat and 'for' not thus blaming without also

(C") Being for blaming for getting others to torture the cat

Schroeder's account, however, is not without its own problems. Perhaps most worrying amongst these is the fact that, in order to be able to account for complex

sentences involving both moral and non-moral aspects e.g. 'Either killing is wrong, or my parents lied to me', Schroeder (2008, Ch. 7) has to posit a similar account of non-moral discourse. Thus, on Schroeder's account, the paradigmatically factual claim

(1) The cat is on the mat

must be explained as expressing

(1') Being for proceeding as if the cat is on the mat

But, quite apart from the fact that such an assimilation of the uses of moral and non-moral utterances is implausible on the face of it, (1) *can't* be explained by saying that it expresses (1'), for (1') *presupposes* the grammar of (1) (cf. Hacker 1990b, p. 383). This is a similar problem to that experienced by 'response dependent' accounts of 'secondary property' terms. Such theories try to analyse statements such as

(2) *x* is red

as being equivalent to

(3) *x* is such as to appear red to persons *P* under conditions *C*

But to know what it is for something to 'appear red' *presupposes* an understanding of what it is for something to *be* red. To say that something 'appears red' draws a contrast with its *being* red and such a contrast cannot be understood independently of an understanding of

what it is to be red (Hacker 1987, p. 117; Z §418).

Indeed, even Schroeder himself is less than optimistic about the prospects for his 'bifurcated attitude semantics'. He is not, after all, an adherent of his own theory, but proposed it by way of seeing to what extent an expressivist semantics could be made to work (Schroeder, M. 2008, ch. 12).

Disagreement Semantics

This brings us to what is, perhaps, the most promising non-cognitivist account to date. Alan Gibbard's (2003) 'disagreement semantics' takes as its starting point the Stevensonian notion that there can be 'disagreements in attitude'. For Gibbard, it is such disagreement which accounts for the logical relations between sentences containing moral predicates.

According to Gibbard, when one person believes that *x* is wrong and another believes that *x* is not wrong, they disagree in attitude with one another. Similarly, someone who believes that either *x* is wrong or *y* is wrong disagrees in attitude with someone who believes that neither *x* nor *y* is wrong and someone who believes that if *x* is wrong, then *y* is wrong disagrees with someone who believes that *x* is wrong, but not that *y* is wrong, and so on.

Now, according to the non-cognitivist, to believe something to be wrong is (roughly[103]) to disapprove of it. So, looking again at Geach's argument:

103 Again, this is a simplification of Gibbard's actual position, according to which a moral utterance expresses acceptance of a norm according to which something is permitted or forbidden. (Gibbard 1990, p. 7) Again, though, substituting the attitudes of approval and disapproval for the sake of brevity and clarity does not, I think, affect the argument.

P1 Tormenting the cat is bad

P2 If tormenting the cat is bad, then getting your little brother to torment the cat is bad

C Getting your little brother to torment the cat is bad

To believe P1 is to disapprove of tormenting the cat and to believe P2 is to disagree with disapproving of tormenting the cat, but not of getting one's little brother to do so. Thus, if someone believes both P1 and P2, he is committed to disapproving of C, on pain of disagreeing with himself.

Gibbard's approach can, perhaps, more plausibly be considered to get the right results as far as the validity of arguments goes than can Blackburn's.[104] Moreover, it doesn't appear controversial that this account captures the seriousness of the kind of incoherence involved in accepting the premises of an obviously valid argument whilst denying its conclusion. It isn't even clear what it would *be* for someone to disagree with himself and someone who appeared to express attitudes which clashed in *this* way would certainly be judged not to understand what he was saying.

Although Gibbard's approach seems *prima facie* promising, though, Mark Schroeder argues that there is a major problem. According to Schroeder, whilst Gibbard has an account of what it is to believe P1, namely, to disapprove of tormenting the cat, he has not, in fact, given an account of what it is to believe P2. All Gibbard has said is that P2 serves to express *something* that disagrees with disapproving of tormenting the cat, but not of getting one's little brother to do so. According to Schroeder, Gibbard hasn't said what the 'mental state' of 'disagreeing with disapproval of *x*, but not of *y*' really *is*, nor has he guaranteed that there *is* such a state. That whatever is expressed by these utterances should disagree is not a legitimate assumption for a semantic theory, rather, that a theory has the

104 If, that is, one accepts Gibbard's (2003, pp. 72f) response to Unwin's (2001) negation problem.

consequences that these statements should disagree with one another is one of the criteria according to which the theory is to be judged correct. According to Schroeder, then, Gibbard simply helps himself to the very notions he is supposed to be explaining (Schroeder, M. 2010a, pp. 131ff).

It may seem hard to see the force of Schroeder's objection. Why, after all, *can't* P2 just be said to express 'disagreement with [disapproval of tormenting the cat, but not of getting your little brother to torment the cat]'? Schroeder, though, has a point. We have, I think, a fairly distinct notion of what is involved in disapproving of something, but it isn't so clear exactly what is involved in 'disagreeing with disapproving of something'. Is it, perhaps, to *approve* of it? If so, then, once again, we have a fairly clear notion of what this means, for 'approval of something' is an everyday notion which we use regularly, whereas 'disagreement with disapproval of something' is a philosopher's term of art. It isn't immediately clear, though, that to disagree with disapproving of something *is* to approve of it, or, in fact, *what* exactly it is to disagree with disapproving of something and to this extent, Schroeder may be justified in asking for some further clarification.

Moreover, whilst it may be true that, say, uttering a conditional moral sentence commits one to disagreeing with anyone who has the attitude that would be expressed by the assertion of the antecedent, but not the attitude that would be expressed by the assertion of the consequent, such sentences are not used only to counter competing points of view (or else we should have nothing to counter.) Hence it cannot be said that what is expressed by such a sentence is disagreement. Rather, the only safe way for Gibbard to characterise the sense in which disagreement is involved with the semantics of moral terms is to say, not that moral utterances are used to disagree, or even that they serve to express disagreement, but that they express *something* which 'disagrees with' other attitudes. But it is normally people who disagree with things. It is not at all clear what it is for attitudes to

disagree, except, perhaps, that the two attitudes are mutually exclusive. Hence, it would appear that Gibbard has, indeed, said no more than that in issuing such a statement, one expresses an unspecified attitude that is incompatible with some specified combination of attitudes.

Perhaps the biggest problem, though, with all extant attempts to grapple with the Frege-point, is their rather narrow focus on formal logic and accounting for the validity of moral inferences. As we have already seen in Chapter Three, the use of an expressive word is far more nuanced than the account on offer by the non-cognitivist. The same sentence that may be used to disagree with someone may also be used simply to inform, a conditional may be uttered ironically, so as to have the effect of denying its own antecedent, and so on. Thus, even if disagreement semantics, or something like it, captures *something* about the use of a moral term, it is still far from capturing its *whole* use. Indeed, given the open-ended nature of linguistic competence, it needn't be *possible* to describe, in advance, every *possible* use of a word or sentence. Whilst recent attempts have been made to widen the focus of non-cognitivist responses to the Frege-Geach point to such features of language, even expressivists admit that they have some way to go in providing anything close to a complete picture of the use of moral terms (Sinclair 2009, pp. 144f). Thus, the non-cognitivist still has an explanatory debt to pay regarding the uses of moral terms.

Idealizing Subjectivism

Non-cognitivists, as I have argued, so far have no pat response to the Frege-Geach point, and the drive to provide one continues to dominate the debate regarding an expressive account of moral language at the cost of providing a metaethic which addresses our worries

and confusions about moral language. Might there, then, be a way to take advantage of the ability of non-cognitivism to explain, for example, the intractable nature of moral disagreement and the conceptual connection between moral judgements and attitudes, without taking on its philosophical baggage? I believe that there is such a way, and that it lies in taking a step back in a debate that appears to have moved on so quickly that it bypassed its destination some years ago.

As previously mentioned, the natural starting point for the motivational internalist is simple subjectivism. This account, however, was more or less abandoned towards the beginning of the twentieth century because of two main objections: the disagreement and modal problems. It is these problems, above all others, which drove the bandwagon towards non-cognitivism in the middle part of the last century (Schroeder, M. 2008, ch. 2; Schroeder, M. 2010, ch. 4).

The modal problem essentially points out that, as *sentences,* '*x* is wrong' and 'I disapprove of *x*' do not have the same meaning. The two sentences cannot simply be exchanged for one another in *any* given context without a change of meaning. For example, '*x* is wrong iff *x* is wrong' is tautologous, whereas '*x* is wrong iff I disapprove of *x*' is false.

The modal problem, then, does indeed dispense with *simple* subjectivism. For, whilst simple subjectivism may correctly capture the meaning of a simple atomic use of a sentence containing a moral predicate, it has failed to capture the meaning of a moral *sentence*, leaving open questions about the multifarious uses of moral terms, analogous to those faced by the non-cognitivist.[105]

It is possible, however, to alter the subjectivist analysis, turning away from 'simple' subjectivism, not towards non-cognitivism, but towards a more 'sophisticated'

105 Indeed, if taken as a claim only about the meaning of a simple moral *utterance*, then there is, by the argument of the previous chapter, no distinction between simple subjectivism and non-cognitivism.

subjectivism. Speaker subjectivists customarily attempt to fix the simple subjectivist analysis by inserting some kind of 'suitable conditions' clause. On such an account, the sentence '*x* is wrong' is said to mean not 'I disapprove of *x*', but something like 'I *would* disapprove of *x* in such-and-such conditions'.

Consider then, the following construal of what is conveyed in passing moral judgement of the form '*x* is wrong':

(W) *x* is such that, after careful consideration, I would disapprove of it.[106]

Such a construal, I think, captures the features of a moral judgement that we are looking for. For, whilst, as I shall shortly argue, it is an expression of disapproval, it *also* allows for doubt and error in our moral judgements, thus avoiding the modal problem.[107]

Whilst, as I shall argue, (W) is properly synonymous with a moral sentence in being usable in the same ways, including in embedded contexts (with, perhaps, one minor caveat, to be added in due course), this sentence 'wears its use on its sleeve', such that it is not in the same way puzzling as that of its synonym. Such an analysis elucidates the use of a moral sentence by making explicit what is only implicitly grasped by those who are competent with the use of the sentence, namely, that a moral judgement makes reference to a *standard* by which one's attitudes are judged correct or incorrect, merited or unmerited.[108]

106 I shall offer an analysis only of the predicate 'wrong', but the account could be applied, *mutatis mutandis*, to other 'thin' moral terms. 'Thick' concepts would require slightly different treatment. This is an area on which much has been written, but into which I shall not stray too far here. Suffice it to say that I take such concepts to be amenable to analysis in terms of the expression of an attitude towards something falling under a particular description. 'Brave', for example, could be analysed as something like 'Good on account of steadfastness in the face of danger or adversity', where 'steadfastness' is analysed purely descriptively. Two people might, therefore, agree that someone showed steadfastness in the face of danger, whilst disagreeing as to whether he was 'brave' (as one might refuse to use the word of an enemy combatant, even though he is enduring the same adversity as one's own troops.)

107 The eagle-eyed reader may suspect that this claim is premature. I shall deal with this worry in due course.

108 I am not claiming that this is what the typical subjectivist takes himself to be doing; his philosophical motivations might be quite different. However, this is, I think, a *virtue* of idealizing subjectivism, even if an unintended one.

The standard is, of course, subjective, as made clear by the indexical element in the analysis, but it is a standard nonetheless, since it allows for confirmation, correction, doubt, error, and, indeed, knowledge.

This allows for a way around the modal problem, for it may be the case that, whilst I now disapprove of something, I would not do so were I to consider the matter more carefully, get more information about the subject, and such like. Hence, on such an account, it is not the case that '*x* is wrong iff I disapprove of *x*'.

Such an analysis, though, is hardly ground-breaking. Numerous attempts have been made to capture moral judgements in such a form, and all have come up against a barrage of objections. Indeed, it may be thought that the analysis does not even overcome subjectivism's central problems with disagreement and modality. In the remainder of this chapter, then, and, indeed, in the remainder of this thesis, whilst I don't pretend to have answered all objections to such a view, I want to address some of the most pressing objections in order to show how such an approach might be rendered at least plausible, on a Wittgensteinian account.

Attitudes and Emotions

The first objection I wish to consider is that such a construal of moral judgements does *not* capture the kind of internalism that we are seeking. For whilst, on such an account, there is *some* kind a conceptual connection between moral judgements and attitudes, a simple assertion of a moral sentence would not, on the above analysis, constitute an expression of disapproval, but would instead constitute only an expression of a *belief* about one's attitudes in some hypothetical situation.

Michael Smith (1986, p. 297),[109] argues that a sincere utterance couched in terms of the attitudes one *would* have in certain hypothetical circumstances entails the hypothetical attitude only on the assumption that the speaker has the additional belief that these hypothetical circumstances are actualized. Put simply 'I would disapprove of *x* in such and such conditions' entails 'I disapprove of *x*' only in conjunction with the additional premise 'I am in said conditions'. Thus a sincere utterance of (W) would entail the speaker's disapproval only on the assumption that the speaker believed that he *had* considered the matter carefully. This would be a bad result, for the internalist must surely allow that people may make instinctive moral judgements in cases where they are not under the illusion that they have thought the matter through carefully, but he nonetheless wants to claim that such judgements constitute expressions of approval or disapproval.

It may seem that an utterance of (W) would at most 'conversationally implicate' that one disapproved of *x* and that this implicature could easily be cancelled by inserting the prefix 'I don't disapprove of *x*, but…', or the suffix '…, but I don't disapprove of *x*'.

This, however, is a little too quick. For it is not at all obvious that this implicature could be thus cancelled. To convey, by way of a speech-act, that one's most considered response to something would be to disapprove of it seems automatically to qualify the speech-act as one of condemnation. Were someone to say 'I would disapprove of *x* after careful consideration' and then add '…but I don't mean by that to deprecate *x* in any way'[110] we would likely take this as some kind of joke. It would be quite absurd to attempt, in the next breath, to retract the obviously disparaging connotation of such a claim. That is, whilst it is always *possible* that a thing which I do not disapprove of is such that, after careful

109 In fact, Smith's arguments relate to a subtly different kind of metaethic – the 'moral sense theory'. The difference between the moral sense theory and idealized forms of subjectivism seems, however, to be more one of metaphysical emphasis than substance. Whilst both proffer similar analyses, the moral sense theorist generally takes moral properties to be metaphysically robust properties that are 'out there' in the world to be perceived, whereas the idealizing subjectivist generally takes moral judgements simply to refer to properties of our attitudes, which are at most 'projected' onto the world.

110 NB – The Oxford English Dictionary defines 'Deprecate' as 'Express disapproval of' (OED online http://oxforddictionaries.com/definition/english/deprecate)

consideration, I would disapprove of it, to say 'This is such that, after careful consideration, I would disapprove of it, but I don't disapprove of it' would be *Moore-Paradoxical*.

I might, of course, deny that my obviously disparaging comment had been intended as such on the grounds that it had not been an appropriate way to convey my true attitude, but this would simply be to retract the comment as inappropriate. Otherwise, to claim that such a derogatory remark as 'I would disapprove of this after careful consideration' had not been intended as an expression of disapproval could only be met with incredulity: 'Well,' one's interlocutor might say, 'I think you made your attitude perfectly clear'.

As W.D. Ross once remarked

> …whatever be true of dislike, it is impossible to disapprove without thinking that what you disapprove is *worthy of disapproval*… (Ross 1939, p. 34)

One needn't think that one *has* scrutinized one's attitudes thoroughly in order to disapprove of it, but one must nonetheless think that one's attitudes *would* stand up to such scrutiny, or else one would have to reserve judgement. This isn't to say that one might not blurt out 'That's wrong' without having thought at all about whether or not one's disapproval could stand up to scrutiny. But, having said such a thing, one would have to admit that one is committed to the claim that one's attitude *would* stand up to scrutiny, or else one would have to retract the statement. The implication, though, also holds the other way around. To think something *worthy* of disapproval, *is* to take it that careful thought about the matter would sustain disapproval. But to say that something is *worthy* of disapproval – that is to say that disapproval is in accordance with the standards for moral judgement – *is* to express disapproval of it.

What lends superficial credibility to Smith's claim is the fact that 'disapproval' and its cognates may appear to connote an *emotional* response, on the model of disgust or delight. This is something of a crucial point, and therefore deserves careful attention.

It seems hard to pin down any *emotional* response that constitutes a necessary or sufficient condition of moral judgement. One needn't, for example, feel *displeased* at what one judges wrong, for one may be unable to refrain from taking pleasure in retribution meted on an enemy, even though one judges it better to turn the other cheek. Nor need thinking one *ought* to do something involve *wanting* to do it, for one may be unable to muster the desire to atone for one's sins, even though one judges that one ought to.

This may seem to fly in the face of internalism, for the internalist claims that having an attitude of approval or disapproval is both a necessary and a sufficient condition of one's holding a moral belief. However, whist there is, indeed, a conceptual connection between emotions, such as pleasure and desire, and attitudes, such as approval and disapproval, the two are of quite different grammatical categories. Emotions are genuine mental states with 'genuine duration', attitudes are more like dispositions, which persist even through unconsciousness, and through various emotional states. Indeed, one's emotional responses are *criteria*, by which others ascribe to one an attitude. A *range* of emotional responses are intelligible as normal manifestations of a particular attitude. This isn't to say that it would be *unintelligible* to experience different emotions in conjunction with a particular attitude. What *is* true, however, is that, given a particular attitude, some emotional responses are abnormal or inappropriate.

It is not possible to state precisely *what* emotional responses are appropriate, given a certain attitude. We must be careful here to separate, on the one hand, what, conceptually speaking, is a normal manifestation of a moral attitude and, on the other, what we find personally acceptable or agreeable. People manifest their moral attitudes in a variety of

ways. For some, the things that strike them as morally imperfect in themselves and in the world around are a source of deep sorrow, even despair. For others, they are a cause of anger. Yet others accept moral failings with remarkable stoicism, perhaps feeling only wistful that the world should not live up to their ideals. Some feel their moral obligations deeply, others take them relatively lightly. To the stoic, the spitfire may seem melodramatic; to the spitfire, the stoic may seem cold. What we find in this way 'appropriate' is a matter of moral personality.

A *family* of emotional reactions are intelligible as normal manifestations of approval and disapproval. Roughly (and trivially), positive reactions are intelligible as normal manifestations of approval and negative reactions are intelligible as normal manifestations of disapproval.

Of course, the situation is vastly complex. Whilst someone may accept, for example, that it is in general appropriate to feel delight at what is good, this rule will admit of exceptions. I may genuinely judge it a good thing that a person should die, rather than live on in suffering, but yet judge that delight would not be the appropriate response to his death, even given these circumstances. After all, it would have been better had he not had to suffer and die at all. Thus I can feel sadness at the loss, but happiness for his peace. Given the circumstances and the facts of human mortality, I may judge what happened good, whilst at the same time judging those circumstances and facts regrettable.

The situation is further complicated by the fact that the strength of emotion that it is appropriate to feel on making a moral judgement will be roughly proportional to the importance one places on the judgement in question. One may, for example, think that someone acted wrongly in breaking a promise, but think it a relatively trivial wrongdoing, meriting displeasure, but not anger. Moreover, one may think an action wrong, and so unjustified, but excusable. Hence even if it is a serious wrong, a strong emotional response

is not called for.

Emotional responses, unlike attitudes, are only defeasibly connected with moral judgement. One may think that something is 'wrong' ('such that, after careful consideration, one would disapprove of it') whilst not having an emotional response normally characteristic of an attitude of disapproval, or *vice versa*, whereas one cannot similarly think that something is 'wrong' ('such that, after careful consideration, one would disapprove of it') whilst not disapproving of it. To think otherwise is to make a category error, for the attitudes of approval and disapproval *involve* thinking a certain range of emotional responses to be 'fitting' or 'appropriate'. That is, it is part of the concept of a moral attitude that emotional responses of a certain kind are normal manifestations thereof.

A passage from Hamlet illustrates the point nicely, I think:

> I have of late, (but wherefore I know not) lost all my mirth, forgone all custom of exercises; and indeed, it goes so heavily with my disposition; that this goodly frame the earth, seems to me a sterrill promontory; this most excellent canopy the air, look you, this brave o'erhanging firmament, this Majesticall roofe, fretted with golden fire: why, it appeares no other thing to me, than a foul and pestilent congregation of vapours. What a piece of work is a man! How noble in reason, how infinite in faculty! In form and moving how express and admirable! In action how like an Angel! In apprehension how like a god! The beauty of the world! The paragon of animals! And yet to me, what is this quintessence of dust? Man delights not me; no, nor Woman neither. (Hamlet, Act 2, Scene 2)

Hamlet judges the world and its inhabitants to be good,[111] but cannot muster any joy or delight at them. He judges his circumstances to be out of the ordinary, putting his lack of positive feelings down to a bout of melancholy. Whilst Hamlet's feelings are entirely negative, his attitude towards the world is not a negative one. He puts his inability to feel delight down to a problem with *himself*, rather than citing features of the world which justify taking a negative attitude towards it and thus make his feelings appropriate. His emotions are taken to be a mere symptom of his melancholy, and are not endorsed as fitting or appropriate, thus showing his underlying attitude to be unchanged by his current state of mind.

Of course, in extreme cases, people's fundamental *attitudes* and not just their current *feelings* can and often do change. In a deeper fit of depression, Hamlet might cease to think the world and its inhabitants good and thus to judge his negative feelings towards them to be inappropriate. He would then come to think of his emotions, not merely as a symptom of mental unrest, but as a merited response to features of the world.

Thus, if I say 'I would disapprove of *x* after careful consideration', then, in adding '…but I don't disapprove of it now', I might be understood as saying that my current *state of mind* is not appropriate to the attitude of disapproval which I clearly manifested in my original statement – that, for example, I very much *want* to do a thing which I judge to be wrong. [112]

Of course, if I *do* then do it, I will be expected to have further emotional responses, such as guilt or shame (cf. LE, p. 7), which act as further criteria of my really having had such an attitude, and therefore having been sincere in my moral judgement. These being

111 I shall assume, for the sake of argument, that Hamlet is being sincere in this passage, an admitted point of contention as far as analysis of the text is concerned. Moreover, I appreciate that Hamlet's judgements are not moral ones, but closer to aesthetic ones. However, like Wittgenstein, I take the two kinds of judgement to be very similar, hence examples from the sphere of aesthetics can be used to demonstrate points about ethics.

112 I might, of course, be understood differently, say, as having somehow had a change of heart after making my original statement.

criteria, the connection is, again, a defeasible one. But again, if I do not have such responses, I will have to admit, if I am going to be taken as having been sincere, that my response is in some way abnormal or inappropriate, and in doing so, I once again express my disapproving attitude. Such discrepancies, though, must be the exception rather than the rule. If a person more often than not had emotions that were out of kilter with his stated moral judgements, then it would become doubtful that he was sincere, or even that he knew what he was saying.

Perhaps, though, there are further counterexamples to the claim that a moral judgement on our analysis would constitute an expression of disapproval. Imagine, for example, that I don't disapprove of something which is disapproved of by a respected role model of mine. In the past, I have always been won round to the same attitude as my role model and I'm therefore confident that I'll come to disapprove of this thing in the end. Can I not, in such a case, say 'I don't disapprove of it, but I would do so after considering the matter carefully'?

It is certainly true that one might, based on the attitudes of those whom one respects, come to *suspect* (even *strongly* suspect) that one would disapprove of something after careful consideration. But one who merely *suspects* that something is the case cannot sincerely say that it *is* the case. If one is completely confident that one would disapprove of something on reflection, then one disapproves of it already, no matter how one *feels* about the matter, for one would then have to admit that there was something awry with one's feelings – that they were merely the result of ill-consideration, rather than a merited response to a form of behaviour (etc.). My sincere assertion that I would disapprove of something after careful consideration would be taken as a manifestation of an attitude of disapproval. If I then add 'but I don't disapprove of it', and this is not somehow taken as an outright retraction of my original claim, it will be taken as an indication that my feelings

towards the thing in question are not appropriate to my long-term attitude. By implication, though, my failure to have the negative feelings appropriate to my long-term attitude is to be dismissed as unbecoming of my considered attitude of disapproval.

Hence, our analysis does not denude moral judgements of their intimate conceptual link with attitudes of approval and disapproval, although it *does* allow for some distance between moral judgements (and therefore attitudes) on the one hand, and specific emotional responses on the other.

Attitudes and Beliefs

It might yet be thought, though, that our subjectivist analysis does not answer the challenge set by the Frege-Geach point. For, in providing such an analysis, we have not given an account of the function of moral predicates in embedded contexts, as the non-cognitivist attempts to do. We haven't said, for example, what is expressed by an utterance of a conditional sentence couched in terms of our analysis.

We might, firstly, question why we should *have* to provide such an explanation. For, unlike the non-cognitivist, we simply say that the *sentence* '*x* is wrong' is used in the same way as the *sentence* '*x* is such that I would disapprove of it after careful consideration'. There is no problem about how this sentence embeds. We understand complex constructions in which this sentence is embedded perfectly well.

A slightly different question might remain, however, about whether our claim that the atomic use of this sentence is to express disapproval coheres with what we know about the sentence's embeddability.

With embedded cases of

(4) I believe that *p*.

this was a simple enough demand to meet. For, as I argued in the previous chapter, self-ascriptions of belief in embedded cases function just as do ascriptions of beliefs to others. It is only in cases of simple atomic use that (4) has a different function, namely to express the disposition that someone else describes me as having when he says '*He* believes that *p*'. Hence to say (4) is to manifest the disposition that is hypothesised in the antecedent of

(5) If I believe that *p*, then *q*.

In manifesting this disposition, I give ample reason to believe that I have it, and therefore to draw the conclusion '*q*' (see Chapter Four).

On our analysis, however, to pass moral judgement by saying

(6) I would disapprove of *x* after careful consideration.

is not[113] to manifest a disposition that is hypothesised in the antecedent of a moral conditional

(7) If *x* is such that I would disapprove of it after careful consideration, then *y*.

For it is essential to our analysis that it is *possible* to be *wrong* about a statement like (6), in a way that it is not possible to be wrong about a statement like (4). But if I can be wrong

113 Or not *always* (see below for further discussion of this point.)

about (6), then it's possible that I'm *not* disposed to disapprove of *x* after careful consideration, in which case (6) is not an expression of a disposition to disapprove of a thing after careful consideration, but a fallible self-*description*. And whilst this self-description is *also* an expression of the disposition of disapproval, this cannot help, since (7) serves not to hypothesize an *actual* disposition on my part, but only to hypothesize my having a disposition in certain hypothetical circumstances, which I needn't believe to be actual. Thus we cannot take advantage of the same explanation of the use of embedded instances of (6) as we gave for embedded instances of (4).

However, the fact that (6) is a fallible self-description (or a description of an action in terms of my attitudes towards it), is not only the problem; it is also the solution. As I have argued, attitudes of moral approval and disapproval, unlike mere likes and dislikes, are essentially taken by those who hold them to be capable of sustaining rational scrutiny. Hence a moral judgement, *pace* the non-cognitivist, is not *purely* an expression of an attitude of moral approval or disapproval, for there can be no such thing.

If the reader finds this incredible, I invite him to consider the fact that there can similarly be no such thing as a *pure* expression of intention. Whilst

(8) I intend to ϕ

and

(9) It's not possible to ϕ

are consistent, it would nonetheless be Moore-Paradoxical to utter both sentences. This is because, whilst (8) is an expression of intention, it is *also* an expression of belief. For the

intention to ϕ *involves* the belief that it's *possible* to ϕ (cf. Schroeder 2006b, pp. 172ff).

Of course, in the case of intention, the implication does not hold the other way around. The belief that it's possible to ϕ does not entail an intention to ϕ. Indeed, not even the belief that one *will* ϕ manages that. However, the belief that it is *not* possible to ϕ does entail that one *doesn't* intend to do so. Hence a belief *can* entail *something* about one's intentions, and if there is such a counterexample to the Humean philosophy of mind, according to which our cognitive and conative natures are *entirely* logically independent, we ought not to be too surprised that there are others. Again, something similar is true of the fact that I believe that my pain was worse yesterday (cf. Chapter Three). For whilst this is a belief, about which I can be mistaken, it entails, strange as it may sound, that I am in pain. It isn't, of course, that the belief causes me to be in pain. Instead, it is simply a grammatical truism that, as a competent speaker, I won't utter these words unless I am in pain, or else want to mislead someone into thinking that I am.

There is, then, on our analysis, no more puzzle about what is expressed by the use of a moral sentence in an unasserted context, such as in the antecedent of a conditional, than there is about what is expressed by a similar use of a prosaic statement of fact. For a moral judgement expresses a *belief* about one's attitudes in the hypothetical circumstances of having considered the matter carefully, and the content of this belief carries over unproblematically into unasserted contexts.

This claim, however, requires some nuancing. For one of the features of (some) moral judgements that we are aiming to capture is their non-contingency. In such cases, whilst my moral judgement will still have the content (W), it would be odd to say that I express a *belief*, and more apt to call this simply an expression of my attitude. For where my attitude is towards a kind of action as such, and not towards the consequences that it may contingently have, I have a kind of first-person authority about my considered

attitudes, and hence my moral judgements. I am not prepared, in such cases, to allow anything to count as falsifying or confirming them. I can give no reason for them, but I will similarly accept nothing as a reason to abandon them.[114]

Such a case, though, is once again like the case of (4) and (5). For an expression of an attitude which is thus unshakeable is, indeed, a manifestation of a disposition to have the attitude after careful consideration. For no consideration could change my attitude, but only, perhaps, a special kind of conversion (of which more in the next section).

The suggestion, however, that a moral judgement expresses an attitude and a belief, and that the content of the belief is what accounts for the entailment and inconsistency relations between moral utterances, may sound like 'hybrid expressivism', and may, therefore seem to be subject to the same objections.

According to hybrid theories, a moral proposition *p* expresses a belief with a content *c*, and an attitude *a*. It is the content, *c*, which does the work in explaining the validity of moral arguments, since moral propositions simply inherit their inconsistency relations from their descriptive content.

Mark Schroeder, however, has argued that hybrid theories have a hard time explaining the inference-licensing status of moral propositions. For, claims Schroeder, if a moral proposition *p* has content *c*, and expresses attitude *a*, one might well accept the premises of a valid moral argument whilst failing to accept its conclusion. All that one would be committed to on coming to accept the premises of a valid moral argument would be a *belief* with the content of the conclusion, and not the attitude which the expressivist claims is expressed by a moral conclusion, since having a belief with content *c* is logically independent of having *a,* and having a belief with the content not-*c* is logically independent of having an attitude which is contrary to *a* (Schroeder, M. 2009, p. 268).

As we can see from the above argument, though, this is not always the case. For, as

114 This will be discussed in more detail shortly.

I have argued, one cannot, despite appearances, believe that one would disapprove of something after careful consideration, yet not disapprove of it, or believe that one would not disapprove of something after careful consideration, yet disapprove of it.

In order to get around this problem, most hybrid theories (and, so Schroeder argues, the most promising ones) model moral utterances on uses of pejorative terms, such as racial slurs. In doing so, they commit themselves to the claim that *all* uses of moral terms express *the same* attitude. For to use a racial slur, even in the antecedent of a conditional, is to express a contemptuous attitude. After all, if I don't want to come across as contemptuous towards a certain race, then my reasoning won't even contain racial slurs – they simply won't be words that I'm prepared to use (although I might be prepared to *mention* them, perhaps in order to explain their meaning, or to tell someone else not to use them). Whilst this may be a plausible account of certain pejorative terms, though, it is a rather implausible account of moral ones. Indeed, it is exactly the view that Geach held up to ridicule in light of the Frege-point (Schroeder, M. 2009, p. 274).

The analysis I have offered, however, does not model itself on such pejorative terms. For one thing, the descriptive content of a pejorative term is the same across speakers, which is not the case with moral judgements on the above analysis, due to the indexical element referring to the speaker's considered attitudes. Mark Schroeder (2009, pp. 284-7) dismisses such an approach essentially on the grounds that it falls prey to the 'disagreement problem'. This, however, is a problem which I shall tackle at some length below.

Moreover, there is no reason why the sentence above should express any attitude when used in an unasserted context, unlike a pejorative term. Hence objections pertaining to such theories pass our analysis by.

More worrisome, perhaps, is the claim that scrutinizing a situation too closely might actually "kill off legitimate valuation" (Johnston 1989, p. 152). There are clearly cases where the acquisition of too much information, or the undertaking of too vivid an imaginative exercise in respect of a question of value would lead to results that we would consider perverse.

If, for example, in order to have a carefully considered view on whether or not it would be wrong to exterminate the human race, I am required to make myself fully aware of the minutest detail of every atrocity ever committed by mankind, every natural disaster that has ever taken place, and such like, it is well possible that such information would send me mad, and that I would come to judge that such a course of action was permissible. Is, then, my considered opinion that of a mad man?

It ought, firstly, to be noted that 'careful consideration' needn't involve 'full information', as many, or perhaps even most (Sobel 2009, p. 337) contemporary 'idealizing subjectivists' take it to. According to the 'full information' account, what is good, say, is not what I now approve of, but that of which I *would* approve were I 'fully informed'.

The notion of being 'fully informed', however, is not as transparent as it may seem. Indeed, as it is employed by many moral philosophers, the notion is patently incoherent. In order to illustrate this point, take a simple example of an action which might come up for moral consideration, say Alan's killing Brian. Ordinarily, in considering whether or not Alan acted wrongly in killing Brian, we might want to know something about Alan and Brian, their relationship to one another, and the circumstances of the killing, perhaps its effects on third parties and such like: Did Alan act in self-defence, or in defence of a third-party? Were Alan and Brian combatants in time of war? Was Alan sanctioned by the Courts

to execute a capital sentence?

By contrast, we will *not* generally consider at what distance the killing took place from Alpha Centauri, or what Kublai Khan would have thought of it. Such considerations are simply *irrelevant* to our moral evaluation of Alan's action. The 'full information' account, however, would have us take knowledge of such complete irrelevancies into account in specifying the 'ideal conditions' of moral deliberation. Indeed to cut an infinitely long story short, the conditions of 'full information' championed by many contemporary idealized subjectivists would include knowledge of facts about the relation of the action to *all* events and states of affairs: past, present, future and, indeed, counterfactual.

Such a characterisation of 'full information', though, is absurd. The point isn't simply an epistemic one – that we can't, as a matter of fact, know everything. That we cannot do so is not simply a matter of the limitations of human abilities to accrue such knowledge (cf. PI §§208, 251). Rather, there is no such thing as this totality of facts. Counterfactuals, for example, are certainly without limit, and statements about the past and future may well be also.

Moreover, we *don't* as a matter of fact, require (*per impossible*) that a person have such knowledge in order that their moral judgements are taken seriously as well-considered, much as we do not, as a matter of fact, require a person to be able to rule out every *possible* source of error before crediting him with knowledge. Our actual standards for such ascriptions are less stringent than the logically watertight ones demanded by the Cartesian sceptic.[115]

Indeed, a form of Cartesian scepticism appears to be a driving force behind the 'full

115 The rejoinder to this position is more often than not to say that we *should* adopt the standards of the sceptic, for only then can error be made a logical impossibility. There are at least two problems with this suggestion. On the one hand, the sceptic is absolutely correct that, for any empirical claim, no non-question-begging, logically sufficient ground can be offered. Thus the concept of 'knowledge' would, on the sceptic's understanding, become quite useless, and we should have to introduce a new term to do the service of our term 'knowledge'. Furthermore, no conclusions about which concepts we *ought* and *ought not* to use can be drawn *a priori. G*rounds can, of course, be given for such a recommendation, but these will not be conceptual, but, for example, pragmatic, or even moral.

information' account. The worry seems to be that one could not *know* that one had all of the *relevant* information unless one had considered *all* the information and ruled out the relevance of certain obscure aspects of an action. This, however, is simply a species of a more general epistemological problem, and not a special problem for the account that I am proposing. The fact is that, in certain circumstances, we will grant that someone *knows* that they have considered the relevant points. It remains a logical possibility that he *doesn't* know, of course, but this is a feature of *all* genuine cases of knowledge.

In responding to our original worry, though, it is significant that the objection can be brought at all. That is, it is significant that I am able to say that certain alterations in my attitudes would be illegitimate. It may be the case that, were I brought face to face with some particularly disturbing fact about human nature, I would lose my mind and become an apologist for genocide, but in my currently lucid state, I judge that this change of heart would be illegitimate. Of course, after I have actually been exposed to this fact, my attitude may change. But not every change of attitude is equal.

If, before I am made aware of a fact, I say that the fact, if true, would not alter the moral situation, my attitude may yet change upon coming face to face with that fact. In most cases, this will simply be because I become aware of some detail that I hadn't previously considered. Take, for example, Hare's infamous 'fanatic', who, as a committed Nazi, thinks that the fact that a person is Jewish is sufficient reason to send him to a concentration camp. He then discovers that he is Jewish himself (Hare 1965, p. 170). If this revelation changes his outlook, then he might be able to cite as a reason some further fact of which he had not previously thought. Perhaps he had been indoctrinated to think of Jews as quite different from himself (cf. RPP I §96; Z §528), and the revelation made him realize that this was not the case. Perhaps, if he had been aware of this fact before, his attitude would have been different. If you had asked him beforehand whether it *would be*

permissible to send Jews to the camps, *were* they just like him in respect, for example of their capacity for suffering, he would have said that it was not.

We needn't, however, be able to give such reasons in *all* cases. Moral judgements are often quite legitimately informed by our emotional responses on coming face to face with a fact. Again, though, the fact that we are able to say in advance that certain changes in our attitudes would be legitimate, and others not, is significant. In some cases, I will say that any experience that changed my mind on a particular matter would be corrupting, as with the question of whether it is permissible to exterminate the whole human race. On many issues, though, I may think it necessary for me to see or do certain things and allow these experiences to have their emotional effect before I am in a position to judge, or I might be quite sure that further experience won't change my outlook, but not be so committed to my moral judgement that I rule out *a priori* that there might be a legitimate emotional influence, or hitherto overlooked reason that might legitimately change my mind. To have one's moral judgement influenced by one's emotional responses in cases where one would always have accepted that one's moral opinion might legitimately be changed in this way is not to be 'converted' in the same way as it is to have one's compassion turned to hard-heartedness by over-exposure to grim facts, when one would have said beforehand that this would not be a legitimate response.

If our fanatic can cite no fact by way of a reason for his change of heart, and it was not based on some emotional influence that he would always have considered legitimate, then we are forced to consider his change of heart to be baldly *caused* by the experience. He is then in the same position as I would be were I to take a drug, or get hit over the head, renounce my lifelong liberalism, and become a staunch authoritarian. In such a case, I come to be in fundamental moral disagreement with my former self. If I now say that I was wrong to have had such wishy-washy liberal views before, I can give no ground for my

claim, and I cannot cite any emotional influence that I would previously have recognised as legitimate; before I took individual freedom to be an overriding concern, now I do not. Thus, when I say that I was wrong, my disagreement with my former self is of quite a different kind than it would have been had I been brought to reject my liberalism on some ground that I would always have recognized, or through some process that I would always have considered legitimate, had I considered it hypothetically.

To be truly committed to a moral point of view is to think that, should one undergo such a radical conversion, one would have been morally corrupted, rather than enlightened, although one knows very well that one's post-conversion self would think just the reverse. Similarly, as an atheist, I would take no experience as legitimate grounds for belief in God, but I also know that I may yet come to have just such a belief after having certain experiences, as do many others. As an atheist I believe that, were I to be thus converted, I would have fallen into error, but I also know that, having been born again, I would think of my current self as having been blind to the fact of God's existence, but knowing this has no more impact on my judgement than does the fact that I know that I might suffer a psychological breakdown and come to adopt beliefs or attitudes that I never could have reached by careful consideration.

This also allows for a response to a further worry. For if we might come to have attitudes that we could never have reached by careful consideration alone, then it may come into question why we should favour our considered judgements as standards of correctness over those which we might reach by other means.

The fact is that we *do* favour our considered judgements. To think morally just is to take one's judgements to be answerable to such a standard. But the fact that we have such concepts is not entirely inexplicable. Radical moral conversions are rare occurrences, and represent such fundamental change to one's personality that it begins to stretch the criteria

of personal identity, to the point where we may say that someone who undergoes such a conversion is not really the same person any more. It should come as no surprise, then, that our moral concepts do not cater for such radical shifts, but instead presuppose that one's personal identity will remain intact.

Is this, though, to rule out, say, divine inspiration as a route to moral enlightenment, and to condemn those who rely on it, and seek to have their eyes opened by revelation, rather than ratiocination, as irrationalists who are incompetent with moral concepts? I think not. For one thing, the religious believer does not simply wait to be struck by a lightning bolt, but usually thinks as long and hard about moral questions as anyone else (often much longer and harder). Usually, when his eyes are opened, this is by recognition of some feature of the situation of which he was not previously aware, or by some emotional influence that he would always have considered legitimate, as it is with the rest of us.

The difference between the religious thinker and the non-believer is that the religious thinker accepts certain things as reasons for a moral conclusion which the non-believer does not. He may literally believe that he hears the voice of God, and that the fact that this voice tells him that a certain course of action is the right one is sufficient reason to take it. Less dramatically, he may allow himself to be struck by a Bible verse which he reads at an opportune time, and which seems apt to the situation, thinking of this as of divine provenance.

Neither are such experiences the preserve of the religious thinker. The atheist, too, may be open to 'signs' that he ought to act, or to judge a certain way – experiences which help him to make up his mind about a difficult moral question. The only difference here is that the atheist does not see these signs as messages from God.

Escaping the 'Evaluative Circle'

Yet another well-worn problem with the kind of view I am proposing is the claim, often made by non-cognitivists in favour of their alternative approach (see e.g. Gibbard 2003, p. 6), that all attempts to analyse moral judgements in non-moral terms are doomed, since they either fail accurately to capture the meanings of their respective *analysanda*, or else covertly employ evaluative notions as much in need of analysis as the moral terms they were introduced to explain.

Even if our analysis is admitted to be correct, then, it may be argued that it fails in its aim of clarifying our puzzling moral concepts, because the term 'careful consideration' is covertly evaluative. Thus we have not escaped the 'evaluative circle'.

The claim that the notion of 'careful consideration' is as philosophically perplexing, or perplexing in the same way, as moral terms such as 'right' and wrong', however, seems fairly implausible. Whilst the standards of careful consideration may be difficult to describe, the kinds of intractable disagreements we encounter when it comes to values do not arise.

Indeed, it is fairly easy to point to aspects of the standard for critical evaluation of our attitudes which it would be obviously absurd to call into question. To deny, for example, the requirement that one's considered attitudes must be consistent, would certainly betoken confusion. Where one wishes to make a moral distinction between two cases, or *kinds* of case, one must also be able to point to a non-moral difference which one takes to justify the distinction. In philosophical jargon, one's moral judgements must satisfy the requirement of supervenience. Someone who didn't take such considerations into account in his moral reasoning would certainly be reasoning poorly, if he was reasoning at all.

That we can always find *some* difference between cases, however, does not open the floodgates to numerous cases of special pleading, for moral judgements must be universalizable (cf. Hare 1965, esp. ch. 2). I might make a highly specific exception to a general moral rule, but I must endorse the same exception made by others in similar circumstances. And if, again, I pick on some circumstance highly specific to the case at hand in order to justify my exception, I must, yet again, be prepared to accept similar judgements made by others. Indeed, the more I try to narrow down a moral judgement so that it fits only the particular circumstances with which I am now confronted, the more general rules I end up accepting which I have to allow others, too, to be guided by in their moral decision-making. Hence the fact that one can pretty much always satisfy the requirement of supervenience by finding *some* feature unique to one's own circumstance is of little help in satisfying the requirement of consistency, since one must be able sincerely to say that one would allow others to count this as a justifying reason in similar circumstances.[116]

Of course, nothing logically rules out one's being able sincerely to make some highly specific moral judgements. In practice, however, we are seldom able to persuade ourselves that we would really be happy for others to act for such flimsy reasons, and that we are not simply trying to make things easy on ourselves.

Of course, consistency is not the *only* requirement of a well-considered moral view. I may, after all, manage to remain quite consistent in my attitudes by remaining blissfully ignorant of the true nature of the actions under consideration. Thus, in order to have considered a matter carefully, a certain amount of knowledge about the subject matter is required. I hope that the argument of the next chapter will help to indicate at least one

116 Indeed, one might go so far as to say that the requirement of universalizability just is the requirement of supervenience, plus the caveat that the non-moral difference which justifies the moral difference between two actions can't be that they are performed by different people. In other words. Moral judgements, even when directed towards a particular action, commit one to judgements about that *kind* of action, regardless of who performs it.

reason why there is such a rational requirement. Moreover, as I shall discuss in Chapter Seven, there are yet more constraints on what can intelligibly count as a moral attitude.

Whilst much more could, of course, be said on this issue, a full-blown discussion regarding the standards of careful thought is beyond the scope of this thesis. I hope, however, to have at least made it plausible that such standards, whilst they may well be complex and difficult to describe, and may even bring up their own philosophical puzzles, are not in the same way philosophically puzzling as the subjective standards of moral judgement.

Truth Minimalism and the Disagreement Problem

Even if we can avoid such relatively peripheral problems, though, it may be thought that our analysis fares no better than simple subjectivism against a more central difficulty: the disagreement problem. For, on our analysis, the following exchange:

(9) *x* is wrong.

(10) *x* is not wrong.

is rendered

(9') *x* is such that, after careful consideration, I would disapprove of it.

(10') *x* is not such that after careful consideration, I would disapprove of it.

The participants in the original exchange appear to disagree with one another, whereas with

our analysis in place, there appears to be no disagreement.

This objection, though, conflates 'disagreement' with 'contradiction'. As Charles Stevenson (1937, pp. 26f; 1942, pp. 82f) pointed out, there can be forms of disagreement which do not involve making contradictory factual claims. We can disagree with each other, not just in belief, but in attitude. There is a very real, and often very serious disagreement between someone who disapproves of something, and someone who does not. Indeed, such disagreements are typically much *more* serious than are factual disagreements, which can usually be resolved fairly easily, or at least to a great extent ignored.[117] For this reason, this form of the disagreement problem has been dubbed the 'shallow' disagreement problem, and is generally considered to have been overcome (Schroeder, M. 2008, p. 17).

There remains, however, a so-called 'deep' variant of the disagreement problem. The problem here is that we can disagree with one another's moral judgements by calling them 'true' or 'false' (Schroeder, M. 2008, p. 17). On our analysis, however, this doesn't seem possible. Take the following exchange:

(11) *x* is wrong.

(12) That's not true.

On our analysis, this exchange is rendered

(11') *x* is such that, after careful consideration, I would disapprove of it.

(12') That's not true.

117 Of course, not all factual disagreements are easy to resolve, but to be in factual disagreement with one another, we must at least be able to agree on what *would* settle the matter. Moreover, such disagreements are usually only important when the facts affect our *interests* in some way.

It would appear that (11) and (11') cannot mean the same thing, because to call (11) false is to disagree with a moral claim, whereas to call (11') false is to make a claim about one's interlocutor's psychology. To put it another way, if you say '*x* is wrong' I can disagree with you by saying 'That's not true', whereas if you say 'I disapprove of *x*' and I say 'That's not true', I am disagreeing with something different. In the first case, I disagree with your moral judgement; in the second, I accuse you of insincerely expressing an attitude of disapproval, and issue no moral judgement whatsoever.

Wittgenstein's accounts of meaning and truth, however, sidestep this worry with relative ease. Wittgenstein espoused a 'deflationary', 'minimalist' or 'redundancy' account of truth, which he summed up on a number of occasions as follows:

> 'p' is true = p
>
> 'p' is false = not-p (PI §136; cf. PG, p. 123, RFM, p. 117)[118]

According to this account of truth, to call a proposition 'true' simply amounts to (re-)stating the proposition, and to call a proposition 'false' simply amounts to stating its negation.

It has generally been assumed that the fact that two utterances differ in truth-value entails that they differ in meaning. In fact, though, that we use the word 'true' of one utterance, but not of another is a fact about the use, and hence the meaning of the word 'true' and *not* about the use, and therefore the meaning of whatever is *called* true.

The following exchanges, for example, differ in meaning:

118 This formulation of deflationism is 'unhappy' for a number of reasons. Firstly, as Wittgenstein recognised (PG, p. 123) the quotation marks can be dispensed with entirely, as it is not the sentence, or the symbol 'p' that is true or false, but what is said, or the proposition expressed by a particular use of 'p'. However, "p is true = p" is ungrammatical. The schema should instead read "It is true that p = p". (Baker & Hacker 1980b, p. 291) Moreover, the schema is not, as Wittgenstein seems to have thought, a *complete* account of the use of the words 'true' and 'false', since it leaves out reference to the use of these words in assenting to or denying something *already said* (Glock 2015, p. 116), and, indeed, the fact that certain syntactic changes must be made when reporting *what* it is that is being called 'true', in order to account for indexical elements in the proposition referred to.

(A) **Charles:** Killing animals for food is wrong.

George: That's not true.

(B) **Charles:** Killing animals for food is such that, after careful consideration, I would disapprove of it.

George: That's not true.

But this is a consequence of the fact that the meaning of the token utterances 'That's true!' and 'That's false!', as with any other utterance, depends on the context of the utterance. The context of George's utterance differs between the cases inasmuch as Charles uses different words in order to express his disapproval.

The words 'true' and 'false' are employed in conversation so as to have the effect of re-asserting (or asserting the negation of) something that has been asserted by one's interlocutor, *after having transposed what was said into reported speech*. Our conversations could, after all, continue:

(A) **Charles:** Killing animals for food is wrong.

George: That's not true.

Charles: What's not true?

George: That Killing animals for food is wrong.

Or, with the substitution of our *analysans:*

(B) **Charles:** Killing animals for food is such that, after careful consideration, I would disapprove of it.

George: That's not true.

Charles: What's not true?

George: That Killing animals for food is such that, after careful consideration, you would disapprove of it.

In spelling out what he is calling 'true', George transposes what Charles said into reported speech, thus changing the explicit indexical elements in the sentence. Thus, someone may make the same assertion in two ways – one containing an *explicit* indexical element, another not – and, based purely on this *syntactic* difference between the two assertions, by convention, the same response 'That's true!' or 'That's false' will mean something different in each case, even though the original assertion had the same meaning in each of its different phrasings. Thus, Charles and George's original exchanges come down to:

(A) **Charles:** Killing animals for food is such that, after careful consideration, I would disapprove of it.

George: Killing animals for food is not such that, after careful consideration, I would disapprove of it.

and

(B) **Charles:** Killing animals for food is such that, after careful consideration, I would disapprove of it.

George: Killing animals for food is not such that, after careful consideration, you would disapprove of it.

respectively. Here, it is clear that it is what George said that had a different meaning in each case, and not what Charles said.

Truth minimalism also allows the form of subjectivism which I am defending to escape a charge of relativism, at least inasmuch as this is usually defined to be the thesis that the truth-value of a moral claim is relative to its use by a particular speaker, or its context of utterance within a particular society. Relativism of this kind is, as Wittgenstein argued (Rhees 1965, p. 24), untenable, because, since to call a moral claim 'true' or 'false' is to re-state it, or to state its negation, and to make a moral statement (negated or not) is to express an attitude towards something, then to call a moral proposition 'true' or 'false' is to express one's own attitude towards a thing. That is, to call a moral proposition 'true' or 'false' is to endorse or reject it. But relativism would appear to suggest that we are bound to call 'true' moral claims which we do not endorse, or 'false' moral claims that we do not reject, simply on the grounds that they do or do not accurately reflect the values of a particular speaker or society.

The standard relativist response is to say that moral claims are 'true-relative-to-system-of-values-*s*', or some such thing, but we needn't say this, nor, as Wittgenstein pointed out (Rhees 1965, p. 24) is it at all clear what it means, unless it means that 'those who adhere to system of values *s* believe the moral claim "*p*" to be true', which is something that all (with the possible exception of those who deny the very possibility of divergent moral viewpoints) can agree to.

On our analysis, by contrast, to call a moral utterance 'true' is simply to repeat the utterance in reported speech. Thus we are not bound to call 'true' moral propositions with which we disagree. Of course, were we to substitute our *analysans* for a moral *analysandum*, we may do precisely that, but then, in doing so, we would not be passing moral judgement, but commenting on our interlocutor's psychology.

Perhaps, though, there is a problem with this solution. For something similar was attempted in the past by Charles Stevenson in response to Moore's original exposition of the disagreement problem in *Ethics* (Moore 1912, ch. 3), wherein Moore brings, in essence, the following argument against simple subjectivism:

(P1) According to subjectivism:

(i) '*x* is right' = 'I approve of *x*'

(ii) '*x* is wrong' = 'I disapprove of *x*'

∴ (C1) If person A approves of *x* and person B disapproves of *x*, then, on the above analyses, they can say, respectively '*x* is right' and '*x* is wrong' and both be telling the truth.

∴ (C2) On the above analyses, *x* may be both right and wrong.

(P2) An action cannot be both right and wrong.

∴ (C3) (i) and (ii) are not correct analyses of the moral terms 'right' and 'wrong'[119]

Stevenson points out that (C2), far from being entailed by simple subjectivism, is, in fact, inconsistent with it. Moore's conclusion (C2), Stevenson argues, is the result of being insufficiently careful to apply the subjectivist analyses proffered in (P1). On these analyses, it *cannot* be the case that *x* is both right and wrong, as this would take the form 'I

119 Paraphrased from Stevenson's reconstruction of Moore's argument (Stevenson 1942, pp. 71-4).

both approve and disapprove of *x*', which is a contradiction (Stevenson 1942, p. 74).[120]

Stevenson locates Moore's fallacy in the transition from (C1) to (C2), noting that, in (C1) the words 'X is right' and 'X is wrong' contain, on the subjectivist analysis, implicit indexical terms referring to A and B respectively, whereas, when these same words appear in (C2), the implicit indexicals contained therein refer to Moore (or whoever is making the argument). Hence Moore's argument contains an equivocation.

Here, however, is where the problem comes in. For Stevenson goes on correctly to note that, therefore:

> It would seem that
>
> (a1) If "X is right," said by A is true, then X is right; and that
>
> (a2) If "X is wrong," said by B is true, then X is wrong (Stevenson 1942, p. 76)

Stevenson, though, denies that subjectivist analyses allow one to accept either (a1) or (a2) on the grounds that, by these definitions

120 It is sometimes claimed that such propositions do not represent contradictions. Stevenson, however, anticipates the response that, in ordinary language, we might say such a thing as 'I both approve and disapprove of *x*', but quite rightly points out that such an utterance is meaningful only when used in the sense 'I approve of certain aspects of *x*, and disapprove of certain other aspects of *x*' (Stevenson 1942, pp. 74f).

To take a commonly cited example, one may, when dieting, say such a thing as 'I both want a slice of cake and don't want one'. What is meant here can be clarified in various ways. One might say 'Well, I do want cake, but I don't want to put on weight'. In such a situation, one's problem is precisely that one *does* want cake. Alternatively, one might say 'Even though cake would be nice, I don't want any, because I'm trying to lose weight'. That is, one *doesn't* want cake *because* it will make one put on weight. Both are simply polite ways of refusing cake, each with, perhaps, a slightly different emphasis.

If, however, on being asked whether one wanted cake, one were simply to reply 'Yes, I want cake, and no, I do not want cake', then we should not know what to do with such a response. If someone were to refuse to supplement this response with some clarification, we should be within our rights to say that he had not answered the question at all, even to the point where it becomes unclear that the respondent knew what he was saying. We might, of course, take such a statement as a sign of ambivalence. In that case, however, we might equally do so with 'It's both right and wrong'.

(a1) is like: If “I approve of X” said by A is true, then I approve of X. And…

...(a2) is like: If “I disapprove of X” said by B is true, then I disapprove of X (Stevenson 1942, p. 76)

Neither of which is true. This is a problem because it would be quite absurd to deny that the truth of someone's moral claim entails that his moral claim is true!

What Stevenson fails to realise, however, is that, given that the truth predicate functions as I have described

(13) If “X is right” said by A is true, then X is right.

is *not* equivalent to

(13') If “I approve of X” said by A is true, then I approve of X.

, for the clause '“I approve of X” said by A is true' in (13') is equivalent to

(14) *A* approves of X.

, whereas the clause '“X is right” said by A is true' in (13) is equivalent to

(15) X is right.

, which is equivalent to

(15') *I* approve of *x*.

Stevenson, then, fails to recognise that whether we call a statement 'true' or not (and what it *means* to call a statement 'true' or 'false' (or to hypothesise its being true)) is importantly dependent not only on the statement's meaning, but also on its syntax, and that the fact that a person's utterances 'I approve of *x*' (or 'I would approve of *x*...') and '*x* is right' differ in truth-value does not show that his utterances differ in meaning, but is merely a reflection of the sensitivity of the use of the truth-predicate to syntactic differences between the utterances being called 'true'.

Anti-Anti-Anti-Realism

The employment of Wittgenstein's 'deflationary' conception of truth, though, raises a spectre. For this conception is more often associated in Wittgenstein scholarship with a very different kind of metaethic. It is worth, therefore, taking a slight detour in order to assess whether the deflationary conception of truth best supports our subjectivism, or the 'anti-anti-realism' prevalent in Wittgensteinian circles.

Perhaps the most explicit advocate of the view in question is Sabina Lovibond in her *Realism and Imagination in Ethics* (Lovibond 1983). Here, Lovibond expounds “a form of moral realism derived from the later philosophy of Wittgenstein” (Lovibond 1983, p. 25), claiming that

> What Wittgenstein offers us, in the *Philosophical Investigations* and elsewhere in his later work, is a homogeneous or 'seamless' conception of language. It is a conception free from the invidious comparisons between different regions of discourse, or (relatedly) between different aspects of mental activity…On this view, the only legitimate role for the idea of 'reality' is that in which it is coordinated with (or, as Wittgenstein might have said – cf. *PI* I §136 – 'belongs with') the metaphysically neutral idea of 'talking about something'. (Thus OC §66: 'I make assertions about reality.') It follows that 'reference to an objective reality' cannot intelligibly be set up as a target which some propositions – or rather, some utterances couched in the indicative mood – may hit, while others fall short – If something has the grammatical form of a proposition, then it *is* a proposition…The only way, then, in which an indicative statement can fail to describe reality is by *not being true* – i.e. by virtue of reality not being as the statement declares it to be…Thus Wittgenstein's view of language confirms us…in the pre-reflective habit of treating as 'descriptive' or 'fact-stating', all sentences which qualify by grammatical standards as propositions. Instead of confining the descriptive function to those parts of language that deal with a natural-scientific subject matter, it allows that function to pervade all regions of discourse irrespective of content. (Lovibond 1983, pp. 25ff)

The 'seamless' conception of language offered us by Lovibond's Wittgenstein is initially attractive, as it appears to offer us a way to steer a course between competing philosophical theories by dispensing with the need for metaphysical non-natural properties corresponding to our concepts in ethics, whilst avoiding both the reduction of moral

properties to properties 'discoverable by natural science' favoured by metaethical naturalists and the alleged eliminativism of non-cognitivists.

Whilst Lovibond, like other neo-Wittgensteinians, is nominally a naturalist realist (Lovibond 1983, p. 25), her naturalism purports to make space for intrinsically action-guiding moral properties amongst the inhabitants of the natural world, rather than relegating them to a mystical, metaphysical realm, *a la* Moore and the intuitionists, or doing away with them altogether, as the non-cognitivist proposes.

This kind of 'middle way' approach is undoubtedly true to the spirit of much of Wittgenstein's philosophy. As Peter Hacker notes of Wittgenstein:

> It seems to have been an almost instinctive maxim of his that where philosophical debate has polarized…we should find out what was *agreed* by all parties…and reject that. (Hacker 1990a, p.34)

What seems to be agreed by the parties in the debate between realists (both naturalist and non-naturalist), anti-realists, and error theorists is that there is some kind of problem with counting purportedly action-guiding moral properties amongst the inhabitants of the natural world. Thus non-naturalists banish such properties to a metaphysical sphere, but may retain their action-guiding nature, whilst error theorists think that such properties are too 'queer' to exist, even in this metaphysical realm. Naturalists, on the other hand, typically reject their action-guiding nature and place them amongst the world of scientific facts, and non-cognitivists simply reject the notion that moral language is in the business of talking about properties at all, thus retaining the action-guidingness of moral judgement, and not being committed to 'queer' properties or metaphysics, but are left with some explaining to do regarding what moral judgements do if they don't make claims about

moral properties, and how exactly the logic of moral language works is they don't do so. Lovibond, on the other hand, simply rejects the notion that this is a problem, and insists that the action-guiding nature of moral judgements does nothing to show that these are not prosaic judgements of fact about a property possessed by actions (etc.), since the standard by which we are to decide whether or not an utterance aims to ascribe such a property is simply whether or not the utterance is "couched in the indicative mood" (Lovibond 1983, p. 26).

Whilst, as I say, this approach seems true to much in the *spirit* of Wittgenstein's philosophy, it is far from true to its *letter*. Lovibond's position has been summed up by David Macarthur as "taking indicative language and a deflationary conception of truth to straightforwardly lead to a kind of across-the-board realism" (Macarthur 2010, p. 81 fn1). As Macarthur points out, though, this is far too quick. It is a leitmotif of Wittgenstein's later philosophy that in order to understand the meaning of a word or a sentence, we must look to its *use* (PI §43) and that this cannot simply be 'read off' from its form.

Lovibond's 'seamless' conception of language has been similarly criticised by Simon Blackburn (1990 & 1998) and, more recently, by Hans-Johann Glock (2015). As both Blackburn (1990, passim. & 1998, pp. 161-4) and Glock (2015, pp. 118-122) point out at some length, the exegetical claim that Wittgenstein offers us such a 'seamless' conception of language is rather implausible. Given the sheer number of counterexamples to the minimalist mantra, I shall not go over all of this same ground, but I would like to point out a few examples which have previously been overlooked.

Wittgenstein gave anti-realist or expressivist treatments of a number of areas of discourse. As both Glock and Blackburn point out, Wittgenstein's treatment of certain first-person psychological utterances is replete with counterexamples to minimalism. One aspect of this extended counterexample which has received little attention by Lovibond's

opponents[121] is his discussion of Moore's Paradox, wherein can be found an almost word-for-word refutation of Lovibond's claim that “the only legitimate role for the idea of 'reality' is that in which it is coordinated with...the metaphysically neutral idea of 'talking about something'” (Lovibond 1983, p. 25). In responding to the paradox, Wittgenstein states that:

> The difficulty becomes insurmountable if you think the sentence “I believe...” states something about the state of my mind. If it were so, then Moore's Paradox would have to be reproducible if, instead of saying something about the state of one's own mind, one were making some statement about the state of one's own brain. (RPP I §501)

As previously explained, Wittgenstein's solution to Moore's paradox was, precisely, to show that credal statements were *not* descriptions of a state of mind or a disposition, but *manifestations* thereof. Thus, to say 'I believe…' is not to *talk about* a state of mind, as one might talk about the state of one's brain.

Lovibond's cites OC §66 (“I make assertions about reality”) in support of her claims. *Pace* Lovibond, however, this passage is not a statement of minimalist realism. Rather, Wittgenstein is here talking about empirical knowledge, which has, of course, to do with reality. Indeed, we could even accept that 'I make assertions about reality' is a grammatical proposition, and yet deny that all utterances couched in the indicative mood are appropriately, or non-misleadingly called 'assertions' or 'statements', as Wittgenstein in fact does in respect of certain psychological utterances:

121 Indeed, Macarthur (2010, p. 84) goes so far as to deny that Wittgenstein could possibly mean to say that first-person credal statements do not describe the same states of affairs that the corresponding third-person statements do.

> To call the expression of a sensation a *statement* is misleading because 'testing', 'justification', 'confirmation', 'reinforcement' of the statement are connected with the word “statement” in the language-game. (Z §549)

Wittgenstein similarly rejected realism in mathematics, arguing that mathematical propositions are not (primarily) statements of mathematical facts, but rules and the logical consequences of rules. Again, Wittgenstein went to great pains to explain that 'grammatical propositions' should not be thought of as statements of fact, responsible to reality for their correctness, but as rules.[122]

One passage, though, deserves special attention, both because it is a passage on which Lovibond herself relies, and because a close reading of its context shows the level of perversity involved in taking from it the moral that Lovibond takes.

PI §136 forms part of a wider discussion about the concept of a proposition (§§134-42), which incorporates an implicit criticism of the *Tractatus* notion that “The general form of a proposition is: This is how things stand” (TLP 4.5). The opening remark of this discussion is particularly revealing:

> Let us examine the proposition: “This is how things are.” – How can I say that this is the general form of propositions? – It is first and foremost *itself* a proposition, an English sentence, for it has a subject and a predicate. But how is this sentence applied – that is, in our everyday language? For I got it from there and nowhere else.
>
> We may say, e.g.: “He explained his position to me, said that this was how things were, and that therefore he needed an advance”. So far,

122 This isn't, of course, to say that such propositions cannot also be employed as plain statements of fact, to the effect that such and such a rule is in force. (Schroeder, S. 2009, p. 107)

> then, one can say that that sentence stands for any statement[.] It is employed as a propositional *schema*, but *only* because it has the construction of an English sentence. It would be possible to say instead "such and such is the case", "this is the situation", and so on. It would also be possible here simply to use a letter, a variable, as in symbolic logic. But no one is going to call the letter "p" the general form of propositions. To repeat: "This is how things are" had that position only because it is itself what one calls an English sentence. But though it is a proposition, still it gets employed as a propositional variable. To say that this proposition agrees (or does not agree) with reality would be obvious nonsense. Thus it illustrates the fact that *one* feature of our concept of a proposition is, *sounding like a proposition*. (PI §134)

Wittgenstein, then, makes the following points about the sentence 'This is how things are':

> (a) It is a proposition.
> (b) It has a use (as a propositional variable).
> (c) To say that it agrees (or does not agree) with reality would be obvious nonsense.

From these observations, it would appear immediately to follow that Lovibond's position is directly at odds with what Wittgenstein is attempting to establish in the very passages to which Lovibond appeals in setting out her stall. The proposition 'This is how things are' is meaningful (has a use), but its use is *not* to describe reality, but merely to refer to some

other statement (Baker & Hacker 1980b, p.288).

Lovibond, though, might see this rebuttal as too quick. *No* type-sentence can be assessed for truth (cf. Schroeder, S. 2009, p.103). On a token occasion of use, however, even the sentence 'This is how things are' can be assessed for truth, precisely by considering whether things are as the sentence to which this propositional variable refers claims them to be.

In this case, though, it is still the sentence to which the propositional variable is referring on this particular occasion that is being assessed for truth, for we can't assess the truth of the bare assertion 'This is how things are'. Indeed, on its own, it would not count as an assertion at all, but as a piece of nonsense. Of course, we might answer the question 'Is it true that this is how things are?' in the affirmative, but to do so is, once again, to refer to a proposition already stated.

Moreover, in the very next remark, Wittgenstein says:

> But haven't we got a concept of what a proposition is, of what we take "proposition" to mean? – Yes; just as we also have a concept of what we mean by "game". Asked what a proposition is…we shall give examples and these will include what one may call inductively defined series of propositions. *This* is the kind of way in which we have such a concept as 'proposition'. (Compare the concept of a proposition with the concept of number.) (PI §135)

In this passage, Wittgenstein is alluding to a previous discussion about the nature of the proposition in §§65-88 (Baker & Hacker 1980b, p.152). In §65, Wittgenstein imagines an interlocutor complaining:

> "You take the easy way out! You talk about all sorts of language-games, but have nowhere said what the essence of a language-game, and hence of language, is: what is common to all these activities, and what makes them into language or parts of language. So you let yourself off the very part of the investigation that once gave you yourself most headache, the part about the *general form of the proposition* and of language."

In response, Wittgenstein says:

> And this is true. – Instead of producing something common to all that we call language, I am saying that these phenomena have no one thing in common which makes us use the same word for all, – but that they are *related* to one another in many different ways. And it is because of this relationship, or these relationships, that we call them all "language". I will try to explain this.

Here follows the now-famous exposition of the notion of a 'family resemblance concept', based on the examples of the concepts of 'game' and 'number' (PI §§67f). Wittgenstein points out that there is no one thing in common to all those things that we call a 'game', but that they are held together under one concept by "a complex network of similarities overlapping and criss-crossing: sometimes overall similarities, sometimes similarities of detail" (PI §66).

In §§88-108, Wittgenstein, in an exceptionally intense moment (even by his standards,) lays bare the confusions that had tempted him into his way of thinking about language in the *Tractatus* (Baker & Hacker 1980b, p.191). The focus of the passage is on

our craving for purity and generality in language and philosophy and the difficulty of approaching a philosophical inquiry unburdened by idealizing prejudices, free from the distorting influence of our expectations about language and logic. Notably, Wittgenstein highlights in particular one source of confusion about the uses of words:

> Our investigation is therefore a grammatical one. Such an investigation sheds light on our problem by clearing misunderstandings away. Misunderstandings concerning the use of words, caused, among other things, by certain analogies between the forms of expression in different regions of language. (PI §90)

At this pivotal moment, Wittgenstein puts centre-stage the notion that superficial similarities in our form of expression are a *root cause* of philosophical confusion.

It seems for a moment, as though Wittgenstein has drifted tangentially to his original point, but in the final remark of the section, Wittgenstein comments:

> We see that what we call “sentence” and “language” has not the formal unity that I imagined, but is the family of structures more or less related to one another. (PI §108)

The whole thrust of the more than forty sections preceding is towards the *avoidance* of assimilation. Wittgenstein's criticism of his earlier self is aimed towards the earlier Wittgenstein's tendency to try to see similarities where none exist, or to exaggerate similarities that do exist. The concept 'proposition' is a family resemblance concept (Baker & Hacker 1980b, p.152; PI §65). To say that a proposition is what can be 'true' or 'false', or

to say that 'A proposition says what is the case' is like saying 'Games are all things that can be played'. This is a trivial grammatical reminder, and therefore quite trivially true, but we so far know *nothing* about a game if all we know is 'that it can be played'.

Again and again, Wittgenstein warns against the kind of assimilation which Lovibond sees him as championing. Wittgenstein's attack on 'referentialism' in particular is just one example of his anti-assimilative stance. Thus, having introduced the famous builders' language-game in PI §2, Wittgenstein asks:

> Now what do the words of this language *signify*? – What is supposed to shew what they signify, if not the kind of use they have? And we have already described that. So we are asking for the expression "This word signifies *this*" to be made part of the description. In other words the descripition ought to take the form: "The word…signifies…". (PI §10)

Wittgenstein admits that, of course, one *can* say e.g., that the word 'slab' signifies a slab, in order, for example, to make it clear that it doesn't signify what is called a 'block' in this language game. But, Wittgenstein remarks

> …assimilating the descriptions of the uses of words in this way cannot make the uses themselves any more like one another. For, as we see, they are absolutely unlike. (PI §10)

Wittgenstein famously goes on to compare the diverse functions of words to the handles in a locomotive, which "all [look] alike...Naturally, since they are all supposed to be handled" (PI §12) but which nevertheless serve quite different functions, and to say

> Imagine someone's saying: "*All* tools serve to modify something. Thus the hammer modifies the nail, the saw the shape of the board, and so on." – And what is modified by the rule, the glue-pot, the nails? – "Our knowledge of a thing's length, the temperature of the glue, and the solidity of the box." – Would anything be gained by such an assimilation of expressions? (PI §14)

Driving the point home, Wittgenstein remarks

> When we say: "Every word in language signifies something" we have so far said *nothing whatever*; unless we have explained exactly *what* distinction we wish to make. (It might be, of course, that we wanted to distinguish the words of language (8) from words 'without meaning' such as occur in Lewis Carrol's poems, or words like "Lilliburlero" in songs.) (PI §13)

This kind of assimilation, though, is precisely analogous to Lovibond's notion of "the metaphysically neutral idea of 'talking about something'" (Lovibond 1983, p. 26). If 'Every word signifies something' means 'Every word has a meaning', then, of course, this is trivially true, If, on the other hand, it means 'Every word is correlated with some object', then this is an unhelpful assimilation of words with quite different uses. Similarly, if we say 'Propositions can be true or false', whilst true, we have, as yet, said nothing about the function of a proposition.

One point on which Lovibond is certainly correct, however, as Simon Blackburn (1990, pp. 201f) has pointed out, is that Wittgenstein is not interested in 'invidious' comparisons between different regions of discourse, as a passage chosen by Lovibond in support of her reading makes quite clear:

> When as in this case, we disapprove of the expressions of ordinary language (which are after all performing their office), we have got a picture in our heads which conflicts with the picture of our ordinary way of speaking. Whereas we are tempted to say that our way of speaking does not describe the facts as they really are. As if, for example the proposition "he has pains" could be false in some other way than by that man's *not* having pains. As if the form of expression were saying something false even when the proposition *faute de mieux* asserted something true.
>
> For *this* is what disputes between Idealists, Solipsists and Realists look like. The one party attack the normal form of expression as if they were attacking a statement; the others defend it, as if they were stating facts recognized by every reasonable human being. (PI §402)

However, far from it being the case, as Lovibond claims, that Wittgenstein here condemns the notion that we should not take at face-value the apparently fact-stating form of ethical propositions – that to do so would be to 'attack the normal form of expression' – Wittgenstein's warning is instead directed against quite a specific philosophical mistake, namely, that of dimly recognising a problematic aspect of the grammar of a proposition, and, on the strength of this, concluding that the proposition is *false* (Hacker 1990b, pp. 499f).

Thus, this passage does not speak against the non-cognitivist (or the subjectivist) who recognises that talk of truth and falsity with regard to moral propositions may be misleading, in that it may make us look for something in the world corresponding to our sentences, and leave us confused when we can find no such thing.

What this passage *does* speak against is an error theory. Indeed, but for anachronism, this passage might have been tailor-made as a rebuke to Mackie's claim that, because moral properties would have to be such 'queer' things, quite unlike anything else in the universe, nothing could, *as a matter of fact*, have these properties (Mackie 1977, p. 38).

Mackie rightly recognised that the logical connection between moral judgements and action or motivation showed that they do not serve to describe moral aspects of reality, but was so wed to the notion that this is what they *must aim* to do that he denounced all moral propositions as *false*. Ordinarily, to call a moral proposition *false* is to take a substantive moral position – to express an attitude. Mackie, however, thinks that there is some other way for a moral proposition like 'Torture is wrong' to be false than for torture not to be wrong. For when he says 'It is not true that torture is wrong', he doesn't use this phrase as one would normally use it, to express his lack of disapproval for torture, but instead to express a confused philosophical belief, born out of dissatisfaction with the misleading fact-stating appearance of moral language.

Wittgenstein's rebuke is directed towards those who 'attack the normal form of expression *as if they were attacking a statement*'. One attacks a statement by trying to show that it says something *false*. As Glock (2015, pp. 119f) points out, Wittgenstein's target here is the notion that a 'way of speaking' can be *at fault*, in the sense of saying something *false.* But a 'way of speaking' cannot say something false; only a statement can say something false.

This *isn't*, however, to say that our way of speaking can't be at fault in any other sense. Wittgenstein is often thought of (and sometimes dismissed as) an 'ordinary language philosopher' – someone who thinks that “ordinary language is all right” (BB, p. 28). But far from thinking that our forms of expression are above reproach, Wittgenstein repeatedly identifies them as the *source* of philosophical confusion (RC §27; OC §199; RPP II §§78,

178; RPP I §§292, 984, 1038; PI §§90, 94, 111, 115, 317, 356, 613; Z §§58, 549; BB, pp. 26, 107). Indeed, far from wanting to gloss over differences between different 'areas of discourse', Wittgenstein very explicitly says that ignoring these differences is a source of philosophical error, and even goes so far as to endorse replacing one form of expression with another, less misleading one, at least whilst we are doing philosophy:

> Our investigation is therefore a grammatical one. Such an investigation sheds light on our problem by clearing misunderstandings away. Misunderstandings concerning the use of words, caused, among other things, by certain analogies between the forms of expression in different regions of language. – Some of them can be removed by substituting one form of expression for another; this may be called an "analysis" of our forms of expression, for the process is sometimes like one of taking a thing apart. (PI §90)

Lovibond is quite correct, then, that Wittgenstein does not seek to denigrate any particular area of discourse, as uniformly saying something *incorrect* or *false*,[123] but he often wants to criticise the form of our expressions as philosophically misleading.

Ironically, though, the anti-non-cognitivists have taken one particular cautionary tale regarding the misleading analogies in our language, and used it in defence of their 'seamless' conception of language. The tale in question is that of the truth predicate.

Whilst Wittgenstein did indeed espouse a 'deflationist', 'minimalist', or 'redundancy' account of truth, it is a far cry from such deflationism to Lovibond's realism. For, whilst it is absolutely correct that there is no way for 'p' to be false other than by its not being the case that p, this does not entail that there is no way for 'p' to be false other than by failing

123 At least not on account of its *grammar*. It may well be the case that we are systematically in error about a certain subject, but this is not because of the grammar we use in talking about it, but because of how the world happens to be.

to describe reality, or, indeed, that there is no way that 'p' could fail to describe reality other than by being false (Glock 2015, pp. 119f). Only prior commitment to a substantial notion of truth could get this result, but it was precisely such a substantial notion of truth that the deflationary account was supposed to debunk.

As Glock (2015, pp. 113-120) points out, the deflationary account of truth is meant to dispel the notion that there is a problem as to the nature of truth precisely by showing that truth is not, as it has traditionally been conceived of in philosophy, a relation between a truth-bearer (belief, proposition, thought, etc.) and a truth-maker. Thus, it has been thought that to call a proposition (thought, etc.) 'true' is to impute to it a particular property. The point of the deflationary conception of truth is that to call a proposition 'true' is *not* to impute to it a property, but simply to (re)state the proposition. In expressing a proposition, though, one does not impute to it a property, but rather, one typically (but not necessarily) imputes some property to something *else*.

Thus truth minimalism is itself a counterexample to Lovibond's claims. For to say 'The proposition “p” is true' is certainly a proposition, but its function is *not* to state something about a proposition (cf. PI §134).

Indeed, as Blackburn (1998, pp. 163f) intimates, the moral of the deflationary story is precisely the reverse of that taken from it by Lovibond. Far from settling any ontological questions, Wittgenstein's point is that truth is too thin a notion to carry any philosophical weight at all. And in this he seems quite right, for, if to assert that a proposition is true is simply to assert the proposition, then, as Frank Ramsey correctly noted, whilst there is no longer any philosophical problem as to the nature of truth, a problem remains as to the nature of assertion (Ramsey 1927, pp. 38f; Glock 2015, p. 117). To call a proposition 'true' may simply be to assert it, but this doesn't rule out that to assert 'p' may not be to describe reality, but, say, to express an attitude towards it.

The Modal *Problem*?

Now, whilst we introduced our analysis as a way for the subjectivist to respond to the modal problem, it may be thought that, rather than overcoming the problem, we have simply masked it behind an added layer of sophistication. For our analysis seems to entail that the following is a conceptual truth:

> (16) If *x* weren't such that, after careful consideration, I would disapprove of it, then *x* wouldn't be wrong.

The problem is that (16) contains a subtle ambiguity. On one reading, (16) is unproblematic, for it simply says that if, say, *x* had different features or properties, then my attitude towards it would be different. If, for example, a man's arm instantly grew back when cut off (and this caused him no pain or distress), then it wouldn't be wrong to cut his arm off, but given that things are not like this, it is wrong to cut his arm off. This is a simple case of 'When the facts change, I change my opinions'.

There is another reading, however, which is less amenable to our analysis. On this reading, *x* might cease to be wrong, due, not to a change in the features or properties of *x*, but in my psychology.[124] I might, after all, take a drug or suffer the proverbial 'knock on the head', and my fundamental attitudes might change in such a way that I would no longer disapprove of *x* even after careful consideration.[125] Thus the analysis appears to provide a no less objectionably mind-dependent account of ethics than does simple subjectivism. For I don't want to say that were I to suffer a brain injury and suddenly change my outlook on,

124 There is also a third possible reading, on which (16) simply says that, were my fundamental moral outlook to change, then I would, quite trivially, no longer *judge* these things wrong. But this reading does not, I think, raise any particular problems.

125 The analysis, after all, allows that people who differ on points of principle may consider matters equally carefully, yet not converge in their moral judgements.

say, gender discrimination or capital punishment, such that, were I to consider the matter carefully, I would no longer disapprove of these things, that they would cease to be wrong.

The subjectivist might try to make use of a possible rejoinder to be found in the writings of Simon Blackburn. Blackburn (1981, p. 179) argues that the problem with (18) is that it expresses a moral judgement of an objectionable kind, in allowing that a change in my psychology might be a *reason* to change my attitudes, and hence my moral judgement.

Such a response, however, is not available to the subjectivist. Indeed, the subjectivist is committed to saying that it would be *unintelligible* to take one's attitudes as *reasons* for moral judgements, since he claims that attitudinal self-ascriptions *are* moral judgements.

On a subjectivist account, such a reason would either be vacuous, since it would say no more than that something is wrong is a reason to judge it to be wrong, or else absurd, since it would say that a change in my attitude would be a reason to change my attitude.[126]

Some subjectivists have attempted to tackle the problem by going 'actually rigidified'. According to the actually rigidified speaker subjectivist, our analysis ought to be rendered as:

> (17) *x* is such that I, *as I actually am*, would disapprove of it after careful consideration.

and that (16) should, therefore, be rendered as

126 The unintelligibility of such reasons can also be backed up reference to the aforementioned requirement of universalizability. One could hardly (except in some very special circumstances (see Chapter Seven, fn142)) universalize the claim that one's *own* attitudes should be reasons for *others* to make a particular moral judgement.

(18) If *x* weren't such that I, *as I actually am*, would disapprove of it after careful consideration, then it wouldn't be wrong.

This appears to solve the subjectivist's problem, since it simply excludes the case in which it is me, and not the features of the action which changes.

This strategy, however, appears to struggle with the kinds of moral counterfactuals that we have already thought to be quite innocent. Consider

(19) If people's arms grew back instantly, then it wouldn't be the case that I, *as I actually am*, would disapprove of cutting people's arms off.

The problem is that it isn't at all clear that trying to hold rigid the reference to my *actual* attitudes within the consequent of a conterfactual makes sense.[127] At best, it would seem that this sentence might be rendered something like

(20) In a counterfactual world where people's arms grew back, a person who had the same attitudes that I have in the actual world would not disapprove of cutting people's arms off.

127 Mark Schroeder (2010a, pp. 79f), makes a similar point against an actually rigidified version of simple subjectivism, offering for consideration the following case:

(R) If there was one extra grain of sand in the rings of Saturn but everything else was the same, then I would still think that stealing money is wrong.

(R') If there was one extra grain of sand in the rings of Saturn but everything else was the same, then I would still think that I, *as I actually am*, would disapprove of stealing money.

The mixture of the counterfactual with the actual in (R) leads, according to Schroeder, to fairly obvious issues concerning which world the 'actually' in (R) refers to – the actual actual world, or the counterfactual actual world!

It seems to me, however, that Schroeder's perplexity only affects a subjectivist who wrong-headedly thinks of 'I think that *x* is wrong' as being equivalent to 'I think that I disapprove of *x*', whereas it is, in fact, simply equivalent to 'I disapprove of *x*'.

And whilst such an analysis is not obviously incorrect, it is, I think, at least more confusing than it is enlightening.

There is, however, I think, a simpler way to be had with the modal problem. The problem arises because of the philosophical tendency to think of an analysis as a 'theory' – a linguistic or conceptual theory, perhaps, but a theory, nonetheless – whose 'predictions' are to be tested against actual linguistic 'data'. The subjectivist 'theory' for example, 'predicts' that it should be perfectly acceptable to say that a thing's moral status may co-vary with my attitudes, but when we compare the theory to actual use (or, perhaps, linguistic intuitions, etc., depending upon one's particular philosophical inclinations) the theory comes up short, since this is not something that we want to say. Indeed, it is not even clear that it makes *sense* to say it.

This, however, is the wrong way to look at a philosophical analysis. Instead, we should look at what it is we are seeking from an analysis. As Wittgenstein puts it:

> The *preconceived idea* of crystalline purity can only be removed by turning our whole examination round. (One might say: the axis of reference of our examination must be rotated, but about the fixed point of our real need.) (PI §108; Wittgenstein's emphasis)

I think that we can concede the premise of the modal problem – that there is a difference in use between a sentence expressing a moral judgement and a sentence on the model of our analysis – whilst denying the conclusion that this somehow undermines the analysis. We can admit, as does the non-cognitivist, that perhaps *no analysans* can perfectly capture the *analysandum* that is a moral judgement. There needn't *be* any sentence containing no philosophically puzzling terms that will do the *exact* service of one

which does contain such a term. Our *analysans* may, indeed, not be capable of being put to work in a construction such as (16), at least on one reading, whereas our *analysandum* sits happily (even if vacuously) in the tautology

(21) If *x* weren't wrong, then *x* wouldn't be wrong.

But to reject the analysis on such slender grounds would be to throw the baby out with the bathwater. For all that we need add, in order to be left with a satisfying analysis is the caveat that, whilst the *analysans* has a virtually identical use to the *analysandum*, it cannot be used in a construction such as (16), when this is interpreted as passing comment on a hypothetical change in one's attitudes. That is, specifically *moral* sentences, when couched in the subjunctive, can serve only to consider hypothetical differences in the *object* of the moral judgement, whereas a counterfactual couched in terms of our analysis can *also* be used to consider hypothetical differences in one's attitudes.[128]

With this caveat added, the analysis still clarifies for us the use of a moral sentence. Of course, there may yet be more exceptions to the synonymy of our *analysans* with its moral counterpart which we have not considered, but so long as these remain relatively few, and don't bring up philosophical puzzles, we may simply expand the caveat.

Indeed, just such a caveat is present at the very heart of Wittgenstein's philosophy. Wittgenstein's dictum that 'meaning is use', after all, contains a fairly hefty proviso:

128 It may be thought that this is simply to restate a version of 'actually rigidified' subjectivism. If this is the case, then my way around the modal problem is simply a clarification of what the actually rigidified subjectivist means, hopefully in less confusing terminology, and I do not, therefore, take this to be an objection to my view, but rather, a solution to the actually rigidified subjectivist's problem.

> For a *large* class of cases – though not for all – in which we employ the word "meaning" it can be defined thus: the meaning of a word is its use in the language. (PI §43 (Wittgenstein's emphasis))

For the most part, the constructions 'the meaning of a word/sentence' and 'the use of a word/sentence' can be used interchangeably. There are, however, counterexamples. Some aspects of a word's use are purely syntactical niceties of 'schoolbook grammar', for example, and are not part of the word's meaning. Hence 'use' is a wider concept than 'meaning', even when applied only to words.

That there are counterexamples to any strict claim of logical equivalence between the phrases, though, does nothing to impugn the truth of the claim that meaning is use, when put carefully, and as little to destroy its philosophical usefulness (cf. Schroeder 2006a, pp. 168-81). Similarly, our analysis of moral terms can bear some counterexamples, so long as the analysis pulls its weight in the round.[129]

Adding this caveat also allows us to account for the non-contingency of ethics. For if a moral counterfactual is used only to consider cases in which it is the properties of the action under consideration which differ, and not those in which one's attitudes do so, then an action's moral status is not contingent upon one's attitudes. Moreover, where, one disapproves of a kind of action as such, and not because of any consequences it contingently has, one's moral judgement will no more be a contingent matter than it is a contingent matter that the action-type in question is the kind of action that it is.

Now it may be thought that the addition of such a caveat is objectionably *ad hoc*, and this is an objection with which I shall deal at some length in the next chapter. There

129 It may be thought that to concede counterexamples is to concede too much, and that the addition of such a hedge as I have suggested is entirely *ad hoc*. In the next chapter, I will deal at some length with the charge that idealizing subjectivist analyses are objectionably *ad hoc*, and I hope that the points I make there will serve to ease any remaining unrest the reader feels regarding my proposed response to the modal problem.

are, however, a few things that can be said immediately by way of alleviating such unease. Firstly, whilst the reading of (16) that I have suggested is problematic for the subjectivist is not, I think, nonsensical, it is at least highly unnatural. We naturally make the distinction between a counterfactual difference in one's attitudes, and a counterfactual difference in the properties of an action via the distinction between (16) and its supposedly logically equivalent

> (22) If I weren't such that, after careful consideration, I would disapprove of *x*, then *x* wouldn't be wrong.

So, whilst (16) *can* be used in the sense that is problematic for the subjectivist, excluding this use from our claim of synonymy is to exclude a use which is naturally explained by means of the different sentence (22), and is, perhaps, only given it by those who have been trained in logic.

Moreover, as mentioned previously, radical changes in our moral outlook, not based on reasons or emotional responses to experiences that we would have previously recognised as legitimately informing our moral judgements stretch the criteria of personal identity to the extent that we may want to say that we are not, strictly speaking, the same person any more. Again, although the criteria for personal identity are not hard and fast, thus making it possible to speak of 'my' attitudes in such situations, rather than, say 'his' attitudes, references to *my* attitudes are most naturally taken to exclude such radical breakdowns of personal identity.

The 'actually rigidified' subjectivist, then, perhaps goes too far in saying that it is my carefully considered attitudes 'as I am now' that constitute the standard of correctness for moral judgements, since I may well accept that, in order to be in a position to judge, I

require further experience, and to have my attitudes shaped by my emotional responses to these, but his analysis comes close to capturing the notion of personal identity that is naturally (but perhaps not necessarily) taken to be presupposed in a statement such as (16).

So whilst such a problematic reading of (16) is *possible*, it is not entirely arbitrary to exclude this use of the sentence from our claim of synonymy with a moral sentence.

I hope, then, in this chapter, to have at least made plausible a way of capturing and clarifying some of the puzzling features of moral discourse that seemingly tempted Wittgenstein towards an expressivist treatment, without having to resort to the non-cognitivist's tortured path, by addressing a number of time-honoured objections to subjectivism. In the final two chapters, by contrast, I want to discuss in depth some less well-worn difficulties for such a view regarding which, I believe, Wittgenstein's writings can be of particular assistance.

VI

Why You Should Be Careful What You Wish For

In his 'Why Idealize?', David Enoch (2005) became the latest to join a growing throng calling for subjectivists to justify their now more or less ubiquitous employment of the device he dubs 'idealization' (Enoch 2005, p. 760). Enoch argues that the 'natural' rationale for the idealizing manoeuvre is unavailable to subjectivists, and so challenges them to come up with a suitable alternative.

Whilst his paper hasn't received a great deal of attention, I think Enoch's challenge brings up pertinent questions about the point of philosophical analyses and their standards of correctness. It also brings up the interesting question of why we should be much concerned to engage with moral concepts on the kind of analysis I have outlined. By way of gaining some clarity on these points, then, in this chapter, I aim to provide a response to Enoch's charge that idealized subjectivism is objectionably *ad hoc*.

Enoch's 'Why Idealize?' Challenge

According to David Enoch, subjectivism springs from the 'internalist' notion that moral judgements are inextricably bound up with our conative or affective states. Thus the natural starting point for the subjectivist, claims Enoch, is simple subjectivism – the claim that to pass moral judgement on an action (person, etc.) is to avow an attitude towards it. Simple subjectivism, though, soon succumbs to the objection that we often desire what we later come to recognise as not good, disapprove of actions that, on reflection, do not merit our disapproval, or come to recognise that there are good reasons for action of which we

were previously unaware. The subjectivist is thus left in need of a sticking plaster which will allow him to keep in place something like the internalism to which he aspires, whilst repelling simple subjectivism's septic consequences.

Fortunately a remedy seems near at hand. For usually, where one's moral attitudes shift, this is because one now has more information about what one was judging, has thought more carefully about the matter, etc. Thus it seems that in concocting his analysis, the subjectivist must turn away from *actual* responses and towards *hypothetical* ones. According to the 'idealizing subjectivist', what is good, say, is what I would approve of in certain ideal conditions of full information, full imaginative acquaintance, etc., or perhaps what some hypothetical idealized observer would approve of or recommend.

According to Enoch, however, this dressing soon comes loose, for it is unclear whether, given his internalist starting point, the subjectivist can provide a compelling rationale for his idealizing manoeuvre which makes it appear anything other than an *ad hoc* fix for his ailing simple subjectivism.

There is, after all, according to Enoch, a natural rationale for the idealizing manoeuvre. The natural rationale for idealizing, Enoch claims, is that one's ideal responses more accurately track the moral *truth* than one's unreflective responses. 'Idealizers' often appear to take advantage of just such a rationale, carelessly claiming that one's attitudes in certain idealized conditions are evidence, indeed *infallible* evidence of a thing's value (see e.g. Lewis 1989, p. 121). As Enoch rightly points out, however, such a rationale is clearly inconsistent with the subjectivist's central claim. If, as the subjectivist holds, facts about our responses in hypothetical idealized scenarios *constitute* ethical truths, then these facts cannot be said to *track* the ethical truth, as there are no such independent truths to be tracked.

Enoch brings out the alleged problem thereby posed for the idealizer by way of

analogy with cases where some 'idealization' is required in order to gain a better epistemic position with regard to empirical truths. If I want to know what the time is, then looking at a watch is a good way to find out. Not just any watch will do, though; I need to look at a watch that keeps accurate time. Thus there is reason to make sure, for example, that the batteries in my watch are not flat, that the mechanism is in working order and such like. Similarly, if I want to know which of two people is the taller, then having a look seems a good idea. But not just any look will do; I need to look from the appropriate angle and distance, and when neither person is sporting a Cuban heel. In each case, the rationale for undertaking the relevant epistemic procedure is the natural one: that doing so will help one's beliefs to track the relevant truth.

Enoch brings his point into even sharper focus through the lens of the Euthyphro contrast. When it's 9:43, a good watch will read 9:43 *because* that *is* the time. In ideal conditions, one person will appear taller than the other *because* he *is* taller. It is not the case that it's 9:43 because my watch says so, or that two people are of particular relative heights because they appear to be so, and (this is the crucial point) if this were the case, then there would be *no rationale* for attempting to put oneself in a better position to *find out* these facts. Whatever one's watch read would be the time; whoever appeared taller would be taller (Enoch 2005, pp. 762ff).

Enoch next considers two alternative rationales which the subjectivist might offer for his idealizing manoeuvre. One possibility is that the rationale for idealization in the moral case is simply that it allows the subjectivist to improve upon the obvious extensional inadequacy of simple subjectivism. Isn't the subjectivist, perhaps, entitled to amend his thesis in order to account for the obvious fact that our initial responses are often found to be at odds with the moral judgements that we endorse after sober reflection? Enoch thinks not, for, he claims, any theory whatsoever can be made to fit the facts, given the requisite

ad hoc alterations. There must be reasons other than sheer extensional adequacy for accepting a theory, since there can be infinitely many theories which can claim *this* virtue.

The problem is compounded, according to Enoch, by the fact that the rationale for espousing subjectivism in the first place was not neutral between actual and ideal responses. As noted, subjectivism is usually motivated by some form of internalism – by the desire to account for the connection between an agent's moral judgements and his *actual* conative states, and therefore his actions. If a theory based on actual responses could be made to work, it would surely be preferred by idealizers. This move away from the subjectivist's founding concerns exerts extra pressure on him to provide a more compelling rationale than mere extensional adequacy (Enoch 2005, pp. 766-9).

Perhaps, though, the subjectivist might appeal to our 'actual justificatory practices' in order to provide a rationale. We think, for example, that a person's competence to make judgements of value is contingent upon his being well-informed about that which he is judging. Value claims based on false information, or borne out of ignorance, we think, are not to be taken seriously. The subjectivist, then, may claim that the best explanation for such practices is that what is, say, of value is what we would approve of, desire, etc. given full information about it.

Enoch, however, turns the tables on such an explanation by imagining a religious practice in which the participants take the deliverances of an 'ideal prophet' to be authoritative regarding what is required by God. Is the best explanation of this practice that what it *is* to be required by God is for it to be prescribed by an ideal prophet? No. Those involved in the practice take the prophet's deliverances to be of interest only because they are presumed to be a reliable indicator of God's commands. Similarly, according to Enoch, the idealizer can't rely on the claim that he has the best explanation of our justificatory practices, because another, better, explanation is already on the table, i.e. robust realism.

Our moral justificatory practices are better explained by the supposition that we are trying to get the answers to moral questions *right* – that our moral practice is (at least tacitly) committed to the existence of moral facts distinct from facts about the psychological outcomes of our moral deliberations and that facts about our idealized responses *track*, but do not *constitute* these facts. Hence this rationale too fails (Enoch 2005, pp. 769-78).

Enoch claims that these three rationales are the only ones to be found in the literature and takes this to be strong *prima facie* evidence that these are the only *possible* rationales (Enoch 2005, p. 779). He briefly assesses the possibility that the subjectivist might simply refuse to enter the arena, claiming that his theory is preferable on other grounds (a more parsimonious ontology or a more workable epistemology, for example) but quickly rejects this gambit on the grounds that the implausibly *ad hoc* nature of idealization outweighs any purported gain in plausibility elsewhere (Enoch 2005, pp. 780f). Thus Enoch throws down the gauntlet to idealizers: provide a rationale or perish!

Equivocation, Analysis and Extensional Adequacy

The first thing that I want to note in response to Enoch's challenge is that, throughout his paper, Enoch appears to equivocate between two importantly different senses of a 'rationale for idealizing'. In its first sense, a 'rationale for idealizing' is a rationale for proffering an 'idealizing subjectivist' analysis of value terms. In its second sense, it is a rationale for following a particular epistemic procedure.

The rationale for proffering a particular kind of philosophical analysis of a concept might seem obvious – the rationale is, surely, that the analysis is *correct*. Now, when Enoch addresses the question whether 'idealization' could be legitimized by reference to

our actual justificatory practices, he appears to be talking about a rationale in this first sense and I think that he is quite correct to bring this up as a possible rationale of this kind. Faithfulness to our practice is an essential part of any rationale for providing a philosophical analysis since it is the faithfulness or otherwise of an analysis to our practice that determines whether the analysis is *correct*. The meaning of a word, as Wittgenstein argued, is its use in the language (PI §43). An analysis of a word which failed to be faithful to our normative practices regarding its use would not be a correct analysis of that word. Thus, if Enoch is correct that a more robust realism is clearly a better candidate for faithfulness to our practice, then robust realism wins the day.

It is not a novel point on Enoch's part that robust realism is *prima facie* to be preferred.[130] This thought is usually backed up by reference to our ordinary ways of speaking. We say, for example, that it is not 'up to us' what is right or wrong, we have a proclivity to call moral judgements 'true' and 'false', to think that there is at most one 'right' answer to a moral question.

All of these, however, are what Wittgenstein called 'surface grammatical' features of our moral discourse (PI §644). Our ordinary ways of talking may tempt us, by their superficial formal appearances to conflate areas of discourse which are, on closer inspection, quite distinct. The subjectivist takes himself to have powerful arguments to show that the idea of objective, mind-independent moral truth is ultimately problematic – that such talk does not cohere with what we say in many other cases and leads, eventually, to outright absurdities (cf. PI §464) – whereas a subjectivist account avoids these pitfalls. The point is not, however, that idealized subjectivism provides the 'best explanation' of our practice, as Enoch would have it; it is that it, unlike robust realism, ultimately correctly *describes* and is *consistent with* our practice with value terms.

Enoch's insistence, then, that robust realism has the upper hand so far as

130 See Sinclair 2012 for an enlightening discussion of 'presumptive' arguments in favour of moral realism.

faithfulness to our practice goes simply ignores the body of argument which subjectivists bring to bear in order to show that superficial aspects of our practice with moral terms can be misleading. Thus the jury is at least still out as to who has the upper hand when it comes to faithfulness to our practice.

Enoch's claim that subjectivists 'idealize' in order to gain 'extensional adequacy' is also highly misleading. In saying that a metanormative claim is 'extensionally inadequate', Enoch seems to have in mind that it entails implausible first-order normative consequences. Thus, according to Enoch

> …if I am disposed to value lifelong solitude, then – endorsing something like Lewis's view of values but dropping the idealization – lifelong solitude is of value, even if my valuing it stems merely from my never having thought through what it would really be like. And, of course, lifelong solitude is not of value. (Enoch 2005, p. 766)

But the seemingly troubling first-order normative consequences of simple subjectivism are not the 'idealizer's' primary concern. In making the idealizing manoeuvre, the subjectivist is trying to forestall the troubling *grammatical* consequences of simple subjectivism. We want, for example, to be able to say such things as 'Whilst I disapprove of this, it might nonetheless not be wrong' or 'It's not the case that an action is wrong if and only if I disapprove of it'. And we *don't* want to say that something *was* wrong simply because I disapproved of it, or that it *would be* wrong *were I* to begin disapproving of it.

But idealizing in order to overcome *these* difficulties doesn't seem so *ad hoc*. In doing so, the subjectivist is simply trying to bring his analysis into conformity with our actual linguistic practices and that it accurately reflects correct word usage is simply what

it *is* for an analysis to be correct.

In the previous chapter, I made one move, in particular, in addition to the 'idealization manoeuvre', which may have struck some readers as objectionably *ad hoc*. I suggested that in the face of counterexamples to my proffered analysis, we might simply add a rider that *apart from such cases*, the uses of these sentences were the same. I hope that what I have said so far in this chapter might serve as a rejoinder to the charge that such a move is objectionable. For, whilst the description of a sentence's use as the same as that of another sentence, save in certain respects, may, perhaps, be untidy, it might nevertheless be *correct*.

Correctness, however, is merely a necessary condition of a good philosophical analysis. An analysis may well be correct, but ultimately trivial and unenlightening. That what is (morally) wrong is what one ought (morally) not to do, is a perfectly correct, but thoroughly uninteresting analysis of the term 'wrong'. The same problems which led us to seek an analysis of the word 'wrong' simply reoccur with the word 'ought', and we are no further towards resolving our perplexities regarding either word.

A further rationale for providing a particular philosophical analysis, then, is that it aids us in resolving a philosophical problem. This is the conception of 'analysis' which Wittgenstein espouses in *Philosophical Investigations* (PI §90). In a correct analysis, *analysandum* and *analysans* must be synonymous – we simply replace one form of expression with another – yet analyses should not be trivial tautologies; they should achieve something. What they achieve, though, needn't be the discovery of a new truth; it can simply be a reminder of what, as competent language users, we already know, that is, of the use of words.

Typical idealizing subjectivists do indeed take their analyses to be of use in resolving particular philosophical difficulties. Enoch himself highlights just such a

difficulty which the typical subjectivist claims his analysis helps to resolve, namely the tension between moral discourse's apparently fact-stating form and its conceptual connections with attitudes and action.

Given that the 'idealizing' manoeuvre brings subjectivist accounts closer in line with our actual linguistic practices, whilst maintaining their useful philosophical function, there seems, *pace* Enoch, to be ample rationale for its employment.

Once again, these comments apply also to my claim that we might add riders to any claim of strict synonymy between our proffered *analysans* and its overtly moral *analysandum*. For, so long as such hedges do not impede the philosophical usefulness of our analysis, and also bring our analysis closer to capturing the use of a moral term, what more objection can there be? Of course, if the addition of such hedges proliferates too greatly, such that the analysis becomes too cumbersome to allow us to see clearly the use of the *analysandum*, then the analysis fails even by the relaxed standards I am proposing. This, however, is not yet the case with our proposed analysis.

Why Tell the Time?

Whilst we may have met Enoch's challenge to provide a rationale for the kinds of move the idealizing subjectivist makes in offering his analysis, though, Enoch's aforementioned equivocation appears to bring up a further problem for the idealizer. For Enoch appears to cast doubt on whether there is any point in engaging with moral concepts so analysed.

This brings us to Enoch's 'natural rationale' for idealizing, which is seemingly a rationale in the second sense outlined above. The rationale for ensuring that one's watch is in working order is a rationale for following a particular epistemic procedure, and, of

course, the rationale here is that one wants to know the time – a truth which is independent of the reading of one's watch.

Enoch gets the 'Why Idealize?' ball rolling by appeal to the notion that we undertake certain epistemic procedures in order to ensure that our beliefs track the truth. Without this rationale in place, claims Enoch, there would be no reason to undertake the procedure. If the time depended on the reading of my watch, then there would be no reason to make sure that the batteries were charged or that the mechanism was in working order.

This, however, is a *non-sequitur*. Of course, if the time depended on the *actual* reading of my watch, regardless of the circumstances, then there would be no reason to 'idealize' my watch by making sure it was in working order, etc. If I simply want to know what my watch says, then just having a look at it is a good way to find out. 'Idealizing' it by changing the batteries would indeed be counterproductive with respect to *this* aim.

But mightn't I also want to know what my watch said under certain conditions, under the conditions, say, that it had fully charged batteries and a mechanism which worked? Then I would have reason to 'idealize'. After all, I want to know what my watch says *under these conditions*, and not simply what it happens to say *whatever the conditions*.

Indeed, there is nothing incoherent in the notion that one might have a concept whereby 'the time' *did* in fact depend on the reading of one's watch under certain conditions. Whilst this concept wouldn't be the same as our concept of time (that concept has connections with the position of the sun in the sky and (more recently) with oscillations in the energy states of the caesium atom at 0K) it would be a perfectly usable concept nonetheless.

Indeed, as well as being *useable*, such a concept might even be *useful*. One can think of a potential point for any concept, given appropriate circumstances. Were I a

foppish aristocrat (or a mad hatter) with no particular commitments, but who, for some reason, wanted to take tea at some more or less regular interval not dictated by my fancy alone (or else I should do nothing but take tea), then I might perhaps simply decide that 'tea time' was whenever my watch said twelve o'clock. However, I should not want 'tea time' to be dictated by the time indicated by my watch *whatever* the circumstances, or else my watch might well stop and so I should never take tea again (or else be stuck in perpetual tea time). But so long as the watch is ticking along reasonably well,[131] I am not too fussed about whether it loses or gains a few seconds here or there. Thus I would have a reason to 'idealize' my watch: I want to know when it's 'tea time'.

Similarly, we engage with our actual concept of time because it serves practical purposes. Telling the time allows us to coordinate with other people, to perform scientific experiments and to do countless other things which make our lives easier, or in other ways better. For beings who had no such interests, the concept of 'time' would indeed be useless. There would, for them, be no 'rationale' for doing what is necessary in order to find out the time.

It's hard, then, to know quite what Enoch's complaint is with regard to this kind of rationale. The fact that a certain truth is not independent of the results of a procedure for obtaining it needn't obviate the point of following the procedure when we want to know (for whatever reason[132]) what the result of following the procedure will be.

131 In order for such a concept to be useful, one's watch would have to operate fairly regularly. If 'tea time' is, say, from 12.00-12.30 according to my watch, then it is no good my watch being so disastrously irregular that sometimes this time period is so short that I cannot possibly take tea in this time, or sometimes so long that I am forced to drink all the tea in China. Similarly, our actual concept of time is based on fairly (and more recently *extremely*) regular occurrences. We chose the oscillation of the energy levels of a caesium atom at 0K as the standard by which the second is defined precisely because this phenomenon is so regular. Were the oscillation period of a caesium atom suddenly to start varying wildly, our current concept of time would cease to be useful. The usefulness of concepts is always conditional upon certain facts being as they are (OC §617).

132 Or even for no reason – For some, knowing the time might be an 'end in itself'. For these people, that it helps them to know the time might be reason enough to set their watch accurately. We can imagine this perhaps, as some form of ritual, practised for its own sake – aiming at nothing.

Sobel's Rationale

Perhaps, though, a question remains as to why we should be interested in engaging with moral concepts on an idealized subjectivist account. This brings us neatly to David Sobel (2009), who, to date, is Enoch's sole respondent.[133] The first prong of Sobel's response concentrates on proffering a rationale which, he claims, is to be found implicitly in a great deal of subjectivist literature, but which Enoch overlooks. According to Sobel

> The rationale for granting the idealized agent information and experience is to provide her with a more accurate understanding of what the option she is considering would really be like. (Sobel 2009, p. 343; cf. pp. 337, 343)

As Enoch points out, however, this simply amounts to an insistence that an agent should be allowed 'full non-phenomenological information' as a means to the end of gaining 'full phenomenological information' (Enoch (Unpublished), p. 8). But, Enoch complains, “we have yet to be presented with a rationale for including full information of whatever sort” (Enoch (Unpublished), p. 8). The request for a rationale simply falls through Sobel's suggestion to become a request for a rationale for allowing the idealized agent full phenomenological information.

Sobel anticipates this request, and his response is that the rationale here is that desires formed on the basis of true beliefs and fuller information about their objects are “more genuinely for their objects” (Sobel 2009, p. 347). A desire for X based on false or incomplete information about its object is, according to Sobel “not genuinely for X as it in

133 Whilst Sobel's responses are made primarily in defence of desire-based accounts of well-being or 'good for', which I do not seek to defend, he correctly notes that his responses, if successful, would be sufficient to find Enoch's general critique wanting (Sobel 2009, p. 339; cf. Enoch, Unpublished, p. 3).

fact is" (Sobel 2009, p. 347).[134]

Enoch concedes that there would be nothing *ad hoc* about such a rationale (Enoch (Unpublished), p. 9). However, he questions whether this rationale is, as he claims of the 'natural rationale', unavailable to the subjectivist on his own terms. As Enoch points out, Sobel simply says of desires that are, in this sense, 'more genuinely *for* their object', that they "form a natural class and have a virtue qua desire" (Sobel 2009, p. 347) which justifies the idealizer in privileging them. But, as Enoch rightly points out, we have still not been given any subjectivist-friendly reason why such desires have this virtue, since Sobel's rationale relies on informed desires possessing a virtue which could not itself be explained subjectively on pain of a vicious regress (Enoch (Unpublished), pp. 11f). Thus it would appear that Sobel has failed to meet Enoch's challenge.

Being Careful what You Wish for

The problem with Sobel's rationale, however, is that it is proffered as a rationale of the wrong kind. Sobel is attempting to give a rationale for providing a particular kind of philosophical analysis. But, as Enoch points out, that doing so makes the analysis dependent on desires that have the mysterious virtue of being 'more genuinely for their object' doesn't seem to get us anywhere on this score from a subjectivist point of view.

As a rationale for engaging with moral concepts as the idealizing subjectivist characterizes them, however, Sobel seems on the right track. The thought involved in Sobel's numerous remarks regarding the importance of an agent's having 'full

134 As Enoch rightly points out, Sobel's point is not the "implausible dichotomous one" (Enoch (Unpublished), p. 9) that desires based on false beliefs about their objects are not really desires for their objects (this, after all, would simply be a contradiction); rather "The point is a continuous one: The more informed a desire is, says Sobel, the more genuinely it is *for* the relevant object" (Enoch (Unpublished), p. 9; cf. Sobel 2009, p. 347).

phenomenological information' and about ill-informed desires being 'less genuinely for their objects' is fairly straightforward and can be demonstrated by way of an example. If I desire to read a Dan Brown novel in the mistaken belief that it contains exemplary prose, my desire is not really to read a Dan Brown novel as it in fact is, but to read a Dan Brown novel as I believe it to be. Thus, my desire will not be satisfied by reading an actual Dan Brown novel.

This idea requires some refinement, perhaps, since, of course, it can be said that I did in fact desire to read a Dan Brown novel and that, therefore, logically, my desire will be satisfied by reading one (cf. Sobel 2009, pp. 346F; cf. Enoch, Unpublished, p. 9). However, my *reason* for wanting to read a Dan Brown novel was that (so I believed) it contained exemplary prose. Thus, in one way, the specification of my desire was incorrect, and in another it was correct, but incomplete. My desire might have been more comprehensively spelt out as a desire to read a Dan Brown novel which (so I believe) contains exemplary prose. This more specific desire will not be, or will only partially be satisfied by reading a Dan Brown novel. Furthermore, given that I have a standing desire not to waste hours of my life poring over turgid prose, this desire too will (unbeknownst to me) be thwarted by the satisfaction of my desire (based on the false belief that Dan Brown novels contain exemplary prose) to read a Dan Brown novel. I desired to read a Dan Brown novel as a means to the end of reading some exemplary prose, or rather, my desire to read a Dan Brown novel involved the desire to read excellent prose, given that the presumed fact that Dan Brown novels contain exemplary prose was my reason for wanting to read one.

I have reason to be better informed[135] about the objects of my desires because it will

135 I do not intend to espouse an 'informed desire' account here, as there is, I believe, more to moral decision-making than simply gleaning more information. Eliminating false beliefs and gaining adequate true ones are, indeed, a crucial aspect of moral decision-making, but so too are *choice*, and reflection upon one's emotional responses.

help me more fully to satisfy those desires. Thus I have a use for concepts which are applicable to all and only those things towards which I would take a certain attitude were I to be well-informed about them. One should be careful what one wishes for, because things might not be as one imagines, and the satisfaction of one's wish may ultimately lead to greater disappointment.

Our moral judgements have conceptual connections with how we act, what forms of behaviour we praise or condemn, encourage or avert, appreciate and resent, and the choices which we make between these opposing responses have huge significance in our lives, on the activities we engage in, on the relationships we have with others, and, last but by no means least, how we feel about ourselves. The attitudes of moral approval and disapproval are universalized, and so apply to one's own actions as much as to those of others. Shame and its cognates (guilt, remorse, resentment, etc.) are normal manifestations of self-disapproval and are, by their very nature, unpleasant feelings which we have reason to avoid. Satisfying one's immediate desires, or conforming to one's unreflective attitudes is of little use to a thinking person who can often expect to have to reform these attitudes and desires in light of new information, or more careful thought.

Of course, this means that moral concepts are only of use to those who *do* in fact reflect on their own attitudes and reform them in the light of emerging reasons. For someone who feels no shame, remorse or guilt, who has no conscience, moral concepts will be largely useless, much like colour concepts are of little use to the colour blind.

The rationale, then, for 'idealizing', in this sense, is that, as social creatures and people of conscience, it is useful to us to do so, thus we have a deep need for the moral concepts the idealizing subjectivist describes.

Sobel, Johnston and Alienation

Sobel's second line of attack is to undermine Enoch's self-confessedly pivotal claim that the initial rationale for espousing subjectivism was not neutral between actual and idealized responses.

On this score, Sobel argues that Enoch's claim regarding the founding concerns of subjectivism is a distortion of the typical idealizer's position. Typically, claims Sobel, idealizers aim to provide an account of values (wellbeing, reasons etc.) which "connects up with an agent's nature in a nonalienating way" (Sobel 2009, p. 348). For Enoch's argument to go through, claims Sobel, he has to show that idealizers fail in this pursuit.

This, according to Sobel, is none too easy a task. Take a typical 'idealizer', such as, for example, Williams, who (according to Sobel) claims that what it is for an agent to have an 'internal reason to ϕ' is for there to be a sound deliberative route from his motivational set to the desire to ϕ. An agent, on this account, is not alienated from his reasons. After all, it is, as Sobel notes, "the actual agent's conative set that is engaged when presented with a more accurate understanding of the options" (Sobel 2009, p. 348). Thus there is, by Enoch's own lights, no charge for the idealizer to answer, given that the idealizer is not, as Enoch claims, in the position of having to perform an *ad hoc* bodge job on simple subjectivism.

Enoch, however, responds with a simple counterexample.

> Consider a devoted Muslim, who with all his heart wants to perform the Hajj (pilgrimage to Mecca). Of course, on an idealized subjectivist account, this does not mean that it is good for him to perform the Hajj. Indeed, suppose that his fully informed idealized counterpart would not desire to go to Mecca

> (nor would he advise his actual counterpart to go to Mecca), because – being fully informed – he would realize that there is no god. Insisting that because of this fact going to Mecca is not good for our devoted Muslim, or that he has no reason to perform the Hajj, would be, I take it, an alienating understanding of his wellbeing and reasons if anything is. (Enoch (Unpublished), p. 16)

The purported counterexample, however, is far from convincing. Firstly, if Enoch's parable has any force, it is surely only against idealized subjectivist accounts of 'good for'. The devout Muslim would surely agree that if there is no God, then performing the Hajj is not good (at least, he could no longer claim that it is good for the reason that it is a display of devotion to God. He may, of course, yet think it good to perform the Hajj on other grounds.)

Even against a subjectivist account of 'good for', however, Enoch's counterexample seems fairly flimsy. There are good reasons to think that the idea that one might be 'fully informed' with respect to the existence or otherwise of God (at least in *this* life) is incoherent. The belief in God is a metaphysical one, which the subjectivist can quite consistently (and plausibly) claim to be unverifiable. The standards applicable to questions of empirical fact, and even to ethical questions simply don't apply here. No sense has been given to the notion that one might *know* that God does not exist. After all, how *could* one know?

Perhaps, though, Enoch's point was, in essence, made in more convincing terms many years ago by Mark Johnston against David Lewis. Johnston famously claimed that

> ...in many cases, vivid and complete imaginative awareness may itself kill off legitimate valuation. Harmlessly frivolous activity, such as dressing up in

> unexpected costumes for a philosophy seminar, is a value and so legitimately valued. However it is of the nature of the value in the frivolous that it doesn't bear too much thinking upon, and certainly not very complete or vivid imagining. Represent it to yourself too completely or vividly and you may not be able to resist anticipating the embarrassment that would lead you to disvalue the strikingly frivolous. Yet so long as we are restricting ourselves to the harmless, when it comes to the frivolous the more striking, indeed the more surreal, the better. *Mutatis mutandis* for the erotic. Wouldn't one advise certain restrictions on awareness of the details of erotic goings-on? At least the practical wisdom of seduction is full of ways of clouding consciousness and partly masking reality.
>
> Still another range of examples involves concern for others. Even if one is initially benevolent, complete awareness of the suffering of the mass of sentient beings would be horrifically depressing, and hardness of heart rather than valuing their release might well be the causal upshot. (Johnston 1989, p. 152)

If extremes of reflection can so radically alter our perspective on matters that, in the face of a fuller acquaintance with the facts, compassion can be usurped by hard heartedness, then the subjectivist's picture would appear to be an alienating one indeed.

I already offered a response to Johnston's point in the previous chapter. There is, however, more that can be said in the current context. Johnston's criticism was originally aimed at Lewis's view that the attitudes constituting a standard of correctness for moral judgements are those that would be baldly *caused* by exposure to the facts. This would be an alienating picture indeed, since it would sever the agent's attitudes from what he took to

be good grounds or legitimate emotional influences. On this account, coming to a considered moral attitude would often involve a radical conversion which would dramatically alter the agent's moral personality.[136] As Sobel rightly points out point, however, this is not usually the way that things happen. Usually, one's attitudes change because one already had underlying attitudes towards certain *kinds* of actions and fuller information simply allows one to realise that one was mistaken in thinking that an action was or was not of a kind towards which one has such an attitude. Of course, holding a 'full information' account himself, Sobel cannot ultimately ward off Johnston's criticism in the way that I suggested in the previous chapter. For there are some facts, exposure to which might well alter one's attitudes in a most alienating fashion. That is, some information may not 'engage with an agent's actual conative set', but might, instead, alter it in a way that, given his current 'conative set', the agent would find perverse.

Johnston's point, however, does raise a further question regarding the rationale I have suggested for engaging with moral concepts – that doing so allows one more fully to fulfil one's desires. For Johnston's point would appear to indicate that the opposite may be true in some cases. Sometimes we may have powerful reasons *not* to consider matters too carefully for fear that such consideration might destroy our state of ignorant bliss.

It is certain that we do sometimes find ourselves with powerful motives not to reflect too carefully on the morality of our actions. Many people, in my experience, vaguely suspect that the production of meat involves practices which they could not ultimately condone and thus avoid investigating the matter too carefully for fear that if they did so they would have to give up eating it – an immediate cost which they feel unprepared to bear.

The veil of self-deception, though, is unavoidably thin. Sooner or later, it is likely

136 I don't doubt that such 'conversions' do occur, and are, indeed, 'alienating', in the sense that an agent is then divorced from what he once took to be a reason. But the subjectivist's analysis is not alienating, since it still ties what an agent takes to be of value to his fundamental attitudes as they currently are.

that facts will be brought to one's attention which are impossible to ignore, forcing one to change one's attitudes and therefore to turn one's disapproval back on to one's own self-deceived actions. Actions borne out of ignorance are often excusable. It is often *hard* to find out details regarding the actions which one is judging. But where one has intentionally remained ignorant for one's own ends there is no excuse.[137] Thus there is reason, too, to avoid self-deception. In the end, the subjectivist holds that one has only to answer to one's own conscience. To live with a clear conscience is the ultimate end of the moral person, thus there is reason for him to engage with moral concepts as the idealizing subjectivist takes them to be. On the other hand, if one is not a person of conscience, then one will indeed have little use for moral terms.[138]

The idealizing subjectivist, then, can properly claim that the correctness and philosophical usefulness of his analyses of moral terms is rationale enough for offering them. Such analyses are not *ad hoc* adjustments of simple subjectivism, but are based on observation of our normative linguistic practices, faithfulness to which can be the only grounds for calling an analysis correct.

But there is reason, too, to engage with moral concepts as the idealizing subjectivist construes them. For most of us it is in our nature to take moral attitudes towards our own actions and the actions of others. Such attitudes have far reaching consequences in our lives, thus we have good reason to be careful that our attitudes are not based on false beliefs about their objects and consequently a deep need for the concepts which the idealizing subjectivist describes.

137 This is not to rely on objective moral standards as to what constitutes an excuse. The point is a conceptual one – 'ought' implies 'can'. It is often not within my power to know all of the facts and in such cases I cannot be blamed for acting from ignorance. Where my ignorance is wilful, though, the universally applicable attitude of moral condemnation inevitably falls on my own actions.

138 Not, however *no* use, for such a person might still use such terms to talk about *other* people's moral commitments, in which case, the implicit indexical element in his use of the term will refer to *them*.

VII

Kinship and Understanding

Anti-realist[139] positions in metaethics have long been subject to the criticism that they are too 'liberal' when it comes to placing restrictions on possible objects of moral judgement. Whilst anti-realists do place some restrictions on what kind of attitude might count as a moral one, these restrictions have usually been purely 'formal' – concentrating on features such as 'universalizability', rather than on the kinds of actions towards which such attitudes can intelligibly be directed.

In what follows, I shall argue that, in his *Remarks on Frazer's Golden Bough*, Wittgenstein provides some respite from the disquietude that many feel regarding the anti-realist's 'liberalism'. I shall argue that the understanding of diverse moral viewpoints sought by the opponents of 'liberalism' can be achieved by recognising that such attitudes often share a common root in certain basic human tendencies.

Not all moral responses, however, are rooted in such kinships. This opens up the question of how moral responses which are not thus rooted are to be distinguished from mere preferences, obsessions and such like. I argue, following Wittgenstein on aesthetic judgements, that these kinds of moral responses are characterised by their particular place

139 Whilst it is currently *en vogue* to say that subjectivism is at least minimally realist, I take the form of subjectivism which I have outlined here to have much more in common with its non-cognitivist cousins than with 'robust' realism. It denies, after all, that 'veridical' moral judgements are cognitions of independently existing 'moral properties'. Although it is true that, on such an account, things have the relational 'property' of being 'such that I would disapprove of them...', to say that they have this property is not, on the account I have outlined, to describe something that could properly, or non-misleadingly, be called a property of the action. That an attitudinal self-ascription can be put in the form of a statement about a form of behaviour does nothing to show that a thing has a property in any robust sense in virtue of my having an attitude towards it. To say 'This is such that I disapprove of it' is, after all, equivalent to saying 'I disapprove of this', and the latter certainly does not ascribe a property to an action. Indeed, it doesn't even (except in a trivial and somewhat misleading sense) ascribe a property to me, but is more appropriately called an expression of my attitude. (cf. Z. §551) Whilst 'I would disapprove of this after careful consideration' might seem closer to the ascription of a genuine property, its simple assertive use is just like that of 'I disapprove of this'. Moreover, its method of verification is very different from that of a prosaic statement of fact. Thus, I would count my position as 'anti-realist'. The reader, however, is not obliged to follow me in doing so.

within a culture. Thus the critics of anti-realism are to some extent vindicated: a moral response does indeed require a background, although not the kind of background they may have thought.

Humean 'Liberalism'

David Hume (in)famously argued that moral judgements could not, at bottom, be deliverances of reason (Hume 1739, II.III.III & III.I.I). For Hume, this is (roughly speaking, and dressed up in contemporary jargon) because beliefs, or judgements of fact, unlike moral judgements, do not entail attitudes of approval and disapproval. Hence where the premises of an argument are purely descriptive (expressive of beliefs) and the conclusion contains a moral judgement (an expression of an attitude) one can always quite consistently accept the premises, yet deny the conclusion. Thus, one cannot reason deductively from purely factual premises to a moral conclusion. In order to do this, one's premises must always contain a moral judgement – an expression of attitude – which cannot ultimately be established by argument from further purely factual premises.[140]

According to Hume, an attitude can be rationally criticised only inasmuch as it is based on a false belief. Notoriously, Hume claimed that

> 'Tis not contrary to reason to prefer the destruction of the entire world to the scratching of my finger. (Hume 1739, p. 463)

There are, according to Hume, no attitudes which rational beings as such must or must not

140 I have, of course, claimed that this is not always the case. I still take Hume to be broadly right, however, inasmuch as accepting a description of an action (person, etc.) other than in terms of one's considered attitudes towards it does not involve taking a particular attitude towards that action (etc.).

take towards states of affairs in the world, and hence also no moral judgements that a rational being as such must or must not make.

With only minor amendments, Hume's liberal attitude towards rational constraints on possible moral judgements has been widely shared by anti-realist metaethicists ever since. Such thinkers generally take it that, whilst to make certain moral judgements might be highly idiosyncratic, there is, in principle, no conceptual bar to holding either a positive or negative moral attitude towards *any* kind of action or form of behaviour.

Foot on Moral Beliefs

This liberal view, however, has been criticized, most notably by Philippa Foot, who, in her influential paper 'Moral Beliefs' (Foot 1958), argued that there must be greater conceptual constraint on what could count as a moral judgement than is typically allowed by those in the anti-realist tradition.

As Foot (1958, p. 85) points out, nearly all protagonists in the metaethical debate now place more logical constraints on what might count as a moral attitude than did Hume. Most, for example, agree that 'ought' implies 'can'. Were Jimmy Olson to insist that Clark Kent ought (morally speaking) to scale the offices of the Daily Planet in a single bound in order to rescue Lois Lane, he would either betray his knowledge of Clark's secret identity (and the abilities it entails) or else be considered incompetent with the moral concept 'ought'. One cannot be obliged to do what it is not within one's power to do and thus, whilst one may feel (self-admittedly irrational) resentment towards someone who fails to act when they could not have done otherwise, one cannot hold them morally responsible, on pain of conceptual incompetence.

Similarly, most now recognise that a moral attitude must, as such, be 'universalized'. That is, moral attitudes are, logically, such that they can be directed only towards certain *kinds* of behaviour, rather than a particular action, performed on a particular occasion, or actions performed by a particular person. If I applaud the actions of a friend, whilst looking down my nose at the same kind of action performed by a stranger, this cannot be called a moral attitude, but might at most be described as a manifestation of my fondness for my friend. If I condemn the actions of others, whilst feeling no compunction about performing the same actions myself, then I am not merely a moral hypocrite; my disdain cannot be called 'moral' at all, but might at best be described as a manifestation of self-regard, misanthropy, or similar.[141]

According to Foot (1958, p. 85), these are the only restrictions usually imposed by anti-realists on the intelligibility of a moral judgement. So long as one's attitude of approval or disapproval is suitably universalized and directed towards an action which one believes it is possible for an agent to perform (or refrain from performing), the attitude counts as a moral one. Thus things have not moved on much since Hume, for on such views nearly any *kind* of action could, in principle, be a suitable candidate for moral approbation or disapprobation.

Foot, however, finds this conclusion intolerable. For it allows, for example, that one might take a genuinely moral attitude of approval towards, say the clasping and unclasping of one's hands a certain number of times an hour. After all, one might perfectly well universalize a 'pro-attitude' towards hand-clasping, and the performance of this action is

141 These points require some qualification. On the first point, moral judgements may, of course, be made regarding highly specific cases, such that particular reasons for the judgement are unlikely, in fact, to be applicable in other cases. This, however, is not to the point, for making such a judgement still requires one to judge that *if* these circumstances were to come up again, a similar action would be judged similarly (Hare 1965, pp. 38f).

On the second, a king, for example, might take different moral standards to apply to himself. But he needn't, on that account, be guilty of special pleading, so long as he thinks that what justifies this difference is the fact that he is the king. That is, his judgement is such that, were someone else the king, he too should be given special moral dispensations, on account of his regal status. Such an attitude towards royalty may be objectionable, but it is not unintelligible.

perfectly within the capabilities of most agents.

But, claims Foot, this is absurd! For Foot, our hand-clasper would be like someone who proclaimed that he was 'proud' of the sky (Foot 1958, p. 86). We can, concedes Foot, imagine very special circumstances in which such a claim would make sense. Perhaps, for example, the individual in question has made a huge impact on international environmental policy, which has protected the atmosphere from pollution, or outright destruction. Such an assumption must be made in such a case because, according to Foot, it makes sense to feel pride only where one believes one has some kind of claim to ownership, or has made some kind of achievement in respect of the object of pride (Foot 1958, p 87). As Foot elegantly argues

> ...with no special background there can be no pride, not because no one could psychologically speaking feel pride in such a case, but because whatever he did feel could not logically be pride. (Foot 1958, p. 87)

Similarly, claims Foot, for an attitude logically to be a moral one, it is not enough that it satisfies the minimal constraints imposed by the anti-realist; in addition, a particular background must be in place. For Foot, this background consists of certain beliefs on the part of the moral judge. Specifically, Foot argues that in order for an attitude to be a moral one as opposed to some other kind of disposition (accepting an order, making a resolution, etc. (Foot 1958, p. 86)), it must be founded on beliefs about the relation of the action in question to human well-being (Foot 1958, p. 94). Someone could take a moral attitude towards hand-clasping, claims Foot, only if a background of, say, religious belief is assumed, according to which the welfare of human beings (in this life or the next) was somehow impacted by the performance or non-performance of such actions.

A Clash of Intuitions

In arguing for her position, Foot (1958, p. 86) employs Wittgenstein's 'private language argument' to bat away the notion that psychological predicates, such as 'pride' and 'moral approval', refer to logically private mental objects, which would have the immediate consequence that such predicates might be applicable irrespective of a suitable background of beliefs.

Foot, however, rightly recognises that such considerations on their own do not establish her conclusion, since the more plausible claim that psychological terms are applied on the basis of behavioural criteria is left untouched by such private language considerations.

Foot's argument, however, that ascriptions of moral attitudes require more than simply behavioural criteria, has seemed to many to be weak. As Foot herself admits

> All I can do here is to give an example which may make this suggestion seem implausible...(Foot 1958, p. 92)

Foot's example – the hand-clasper – is supposed to appeal to our intuition that there is something odd about calling such a person's attitude a moral one. Were someone to claim of hand-clasping that it was morally good, claims Foot, we would ask them 'How do you mean?' (Foot 1958, p. 92). Foot argues that such a question can only be satisfactorily answered by giving a certain kind of *reason* for the statement – namely one which connects it with human well-being.

To state this alone, however, is simply to beg the question. For Foot's opponent claims precisely that the meanings of moral terms are not explained by reference to the

kinds of action to which they apply, but by reference to the kinds of attitudes involved in applying them. To say that the meaning can only be explained by giving a certain kind of reason is simply to rule out that an acceptable answer to the question might, for example, be 'I mean that I approve of hand-clasping'.

Of course, Foot's response is that one can't simply state that one has a moral attitude of approval without the correct background of beliefs. She points to the hand-clasper as evidence of this. The strangeness of saying that the hand-clasper has a moral attitude, Foot argues, shows that a distinctively moral attitude must be characterised more restrictively than as a general disposition towards, say, doing a thing, encouraging others to do it, saying 'That's good' when they do, and such like.

But what is it that is really so strange about the hand-clasper? It seems that it is *only* that he doesn't have a certain kind of reason for his attitude. Indeed, Foot's argument *relies* on the fact that the *only* difference between our moral attitudes, and the disposition of the hand-clasper is the kind of reason he takes to justify the disposition.

It seems, then, that Foot and her opponent are simply at loggerheads. Foot refuses to call the hand-clasper's disposition 'moral approval', her opponent sees nothing wrong with such nomenclature.

Wittgenstein on Frazer: The Grounds of Ritual Practice

I think, however, that the deadlock can be broken, and in a way which allows both positions to be given their due. Foot, it seems to me, is right to claim that the puzzlement that most people will feel when they encounter moral views which differ radically from their own is a puzzlement as to *why* anyone would hold such views.

To ask this question, though, is to ask for a *reason* to hold the view, and, of course, if the subjectivist is right, where different moral standards are in play, such a reason will not be forthcoming, and our question remains unanswered – our puzzlement unresolved.

This needn't mean, however, that the understanding sought is entirely unattainable. Wittgenstein's writings on ritual practices, I shall now argue, point to a way in which the satisfaction sought in asking why someone holds a moral view can yet be had, even in cases where fundamental moral agreement cannot be reached. For if we look much more generally at human tendencies, there are ways of understanding another person's actions other than in terms of reasons which *we* accept.

One of Wittgenstein's few pronouncements on matters metaethical is to be found in G.E. Moore's notes from lectures given by Wittgenstein in Cambridge in 1933. Here, Moore recalls that

> He concluded...by a long discussion which he introduced by saying “I have always wanted to say something about the grammar of ethical expressions, or, e.g. of the word 'God '”. But in fact he said very little about the grammar of such words as “God”, and very little also about that of ethical expressions. What he did deal with at length was not Ethics but Aesthetics, saying, however, “Practically everything which I say about 'beautiful' applies in a slightly different way to 'good'”. His discussion of Aesthetics, however, was mingled in a curious way with criticism of assumptions which he said were constantly made by Frazer in the “Golden Bough”... (Moore 1955, p. 16)

What Moore recalls of Wittgenstein's remarks on aesthetics is scant, but nonetheless of some interest for two main reasons. Firstly, they hint at Wittgenstein's

liberal leanings. In particular, Wittgenstein is keen to point out, *a la* Hume, that aesthetic attitudes are not ultimately answerable to reason. Reasons in Aesthetics, claimed Wittgenstein, are

> ...of the nature of further descriptions... if, by giving "reasons" of this sort, you make another person "see what you see" but it still "doesn't appeal to him", that is "an end" of the discussion... (Moore 1955, p. 19)

Here we see Wittgenstein's endorsement of the Humean notion that awareness of facts is insufficient to be moved aesthetically, and thus, by implication, insufficient also to be moved morally. At a certain point, when all the facts are on the table, yet disagreement remains, no further argument can rationally break the deadlock.[142]

Of more interest still, however, are Wittgenstein's 'curious' comments on Frazer's *Golden Bough.* In particular, Moore recalls that, in discussing this topic, Wittgenstein insisted that Frazer had made the mistake of supposing that there is only one reason or motive for performing an action, namely, in order to "get something useful" (Moore 1955, p. 19). This is a view that comes across particularly strongly in the notes published as *Remarks on Frazer's Golden Bough* (RFGB).[143] Here, Wittgenstein insists that ritual practices are characteristically *not* thought by those involved to be efficacious in getting something they want or need:

142 For a more explicit endorsement of this view see, e.g. CV, p. 23

143 Wittgenstein began writing these remarks in 1931, but it would appear that he never substantially changed his mind about the matter, as he continued to add to the remarks until at least 1936, and probably until after 1948 (RFGB, p. 115).

> ...one could begin a book on anthropology by saying: When one examines the life and behaviour of mankind throughout the world, one sees that, except for what might be called animal activities, such as ingestion, etc., etc., etc., men also perform actions which bear a characteristic peculiar to themselves, and these could be called ritualistic actions. (RFGB, p. 129)

Wittgenstein did not mean to deny that people often *do* think their rites efficacious. What he wanted to deny was that such erroneous beliefs were essential to ritual practices as such.

In *The Golden Bough*, Frazer not only describes various ritual practices from around the world, but also attempts to explain *why* people engage in such practices. Frazer's answer, according to Wittgenstein, is that these practices are engaged in because the participants hold false beliefs about the efficacy of the practices. These beliefs, Frazer says, are

> ...not wilful extravagances or the ravings of insanity, but simply hypotheses, justifiable as such at the time when the were propounded, but which a fuller experience has proved to be inadequate. (Frazer 1890, p. 264)

Wittgenstein, however, poured scorn on Frazer's claim. To insist that all ritual practices must be based on such false beliefs, claims Wittgenstein, would be to treat them as 'pieces of stupidity'. But to interpret most ritual actions in this way would be extremely implausible (RFGB, p. 119). In order to demonstrate this, Wittgenstein proceeds to give a number of counterexamples. Some such practices, for example, are clearly intended to be symbolic. The Near-Eastern ritual of drawing an adopted baby through the adoptive

mother's clothes, for example, is a purely symbolic act of re-birth. It would, observes Wittgenstein, be *insane* to imagine that such a ritual is based on the erroneous belief that the mother thereby literally gives birth to the adopted child (RFGB, p. 125). When, on the other hand, a tribesman stabs an effigy of his enemy, he *may* believe that this harms him. However, he *need* not do so in order for the action to be intelligible. Such an action may simply be an expression of hatred, as kissing a picture of one's loved one is an expression of love, in which the picture is used as a substitute, and where such an action is obviously not performed in the belief that it has any effect on the subject of the picture (RFGB, p. 125). Similarly, beating the ground with one's stick when one is angry is not done in the belief that one is thereby punishing the earth; it *expresses* one's anger (RFGB, p. 137).

There need not, then, be only one kind of reason why ritual practices are undertaken. Indeed, there may be no reason at all. A ritual practice, claims Wittgenstein

> ...aims at satisfaction and achieves it. Or rather: it aims at nothing at all; we just behave this way and then we feel satisfied. (RFGB, p. 123)

By implication, then, there need not, according to Wittgenstein, be only one kind of reason for holding a moral view. Foot's insistence, however, that a response cannot be a specifically moral one unless it is a response to a perceived 'human good or harm' is precisely the claim that there is only ever one kind of reason for a moral attitude. It is essentially the claim that moral judgements involve beliefs about the efficacy of certain kinds of action.

Foot takes it that we baulk at the hand-clasper because his attitude towards hand-clasping is not based on any beliefs about the effects of hand-clasping on human well-being. It seems, however, that Foot's claim, like Frazer's claims about ritual practices, is

subject to straightforward counterexamples. We do not have to look far in order to see that we do not, in fact, take moral attitudes to be unintelligible when unaccompanied by thoughts of the impact of actions on human interests. Many people believe that there are circumstances where promises should be kept, inequalities eliminated or punishments doled out even where nobody benefits. It is quite possible for someone to believe that the only harm in performing an action is moral (cf. Mill 1859, pp. 21f) (some people's attitudes towards, say, homosexuality are a prime example) and that the wider 'welfare' of the participants or anybody else is simply irrelevant to the moral status of such actions.

Intelligibility and Kinship

In the opening comment of his *Remarks on Frazer's Golden Bough,* however, Wittgenstein says that

> One must start out with error and convert it into truth.
>
> That is, one must reveal the source of error, otherwise hearing the truth won't do any good. The truth cannot force its way in when something else is occupying its place.
>
> To convince someone of the truth, it is not enough to state it, but rather one must find the *path* from error to truth. (RFGB, p. 119)

Wittgenstein realised that it was not enough simply to give counterexamples to Frazer's claim that ritualistic practices need not be based on false beliefs about their efficacy. For to do this alone will not remove the puzzlement felt with respect to such

practices, making it tempting to reject the truth that ritual practices need not be so founded and retreat to the illusory comfort of Frazer's explanations.

The presentation of purported counterexamples is a similarly inadequate response to Foot's claim that all genuinely moral views are grounded in reasons which make reference to human well-being. For, firstly one may always bite the bullet and deny the counterexamples, claiming that such attitudes *must* in the final analysis, be based on thoughts of efficacy,[144] and that attitudes lacking such grounds are not genuinely moral.

Moreover, whilst it is easy to produce counterexamples to Foot's claim, it is not obvious that doing so addresses her point about the hand-clasper. Whilst Foot may have been proved wrong in her diagnosis of the strangeness of the hand-clasper's case, the strangeness of the case does not seem thereby to be ameliorated.

In this spirit, Wittgenstein intimates that, whilst a ritual practice need not be founded on, or, indeed, involve, beliefs about its efficacy, in order for a ritual to be intelligible to us, it must, nonetheless, be somehow connected to a proclivity which we ourselves share:

> ...if Frazer's explanations did not in the final analysis appeal to a tendency in ourselves, they would not really be explanations. (RFGB, p. 127)

> Once such a phenomenon is brought into connection with an instinct which I myself possess, this is precisely the explanation wished for; that is, the explanation which resolves this particular difficulty. (RFGB, p. 139)

Wittgenstein goes on to suggest that ritual practices are often connected with certain very general human tendencies, which we share with those whose practices we seek

144 Perhaps claiming that sticking to a principle in a case where the benefits of doing so are not obvious is, in the end, based on a belief that holding to principles without exception is somehow beneficial – a claim to which it is not too hard, in turn, to find counterexamples.

to understand:

> It goes without saying that a man's shadow, which looks like him, or his mirror-image, the rain, thunderstorms, the phases of the moon, the changing of the seasons, the way in which animals are similar to and different from one another and in relations to man, the phenomena of death, birth, and sexual life, in short, everything we observe around us year in and year out, interconnected in so many different ways, will play a part in his thinking (his philosophy) and in his practices, or is precisely what we really know and find interesting.
>
> How could fire or the similarity of fire to the sun have failed to make an impression on the awakening mind of man?...(RFGB, pp. 128f)

The idea that there are basic human tendencies common to us all, might seem to suggest that Wittgenstein is here endorsing something like Foot's naturalism – that moral beliefs must have something to do with 'human goods' as determined by our common nature as human beings. Wittgenstein, however, continues:

> I don't mean that just *fire* must make an impression on everyone. Fire no more than any other phenomenon, and one thing will impress this person and another that. For no phenomenon is in itself particularly mysterious, but any of them can become so to us, and the characteristic feature of the awakening mind of man is precisely the fact that a phenomenon comes to have meaning for him. (RFGB, p. 129)

As a preliminary, it is necessary to address what may appear to be an air of tension within these passages. For it appears that Wittgenstein vacillates between, on the one hand, the notion that certain phenomena – those related in various ways to very general facts of human existence – are naturally impressive or important to human beings, and, on the other hand, the liberal notion that no phenomenon is in itself important, indeed that *any* phenomenon *can* be so.

In response to this worry, however, it should first be noted that one can consistently hold that there are, as a matter of fact, certain things that human beings find important by virtue of our common human nature, whilst recognizing that they are not important or meaningful 'in themselves', but only insofar as they relate to the lives of human beings. It is not, after all, part of the concept of fire that it is impressive to human beings, but it is, nonetheless, true that, by and large, it is (RFGB, p. 143).

Furthermore, whilst, as Wittgenstein says, any phenomenon might *become* important, some are more directly and obviously linked to human nature or instinct than others, such that one phenomenon can gain importance *via* the instinctive importance of another.

The clasping of hands is not instinctively important to us, but it could understandably become important. The ritual of clasping and unclasping one's hands would become immediately intelligible were it to be explained that it was symbolic of washing one's hands, even if those involved did not believe that it was in any way efficacious (as a prophylactic against illness, for example). Cleanliness is a more or less universal human instinct (to one degree or another), and it would not be puzzling at all if it turned out that our hand-clasping fanatic saw hand-clasping as obligatory because, for him, the gesture was symbolic of hand-washing. But he need be under no illusion that such symbolic action is in any way efficacious. Indeed, he may recognise that adherence to this practice comes at

a cost, and yet remain steadfast in his adherence without thereby showing himself to be incompetent with moral concepts.

One might even imagine a culture where ritual cleansing was considered *more* important than *actual* cleansing. This becomes especially plausible when we recognise that a cleansing ritual need not even be thought to be celebratory of the importance of cleanliness. Think, for example, of baptism, wherein it is (at least in part[145]) a figurative cleansing of the *soul* that is important (cf. RFGB, p. 125). But cleansing the soul can only be important to us because cleansing our bodies is so. The meaning of the ritual here is metaphorical, but the significance of the metaphor relies on the significance of the object of comparison involved.[146]

It is noteworthy that our understanding of ritual action is itself based on the recognition of a tendency within us towards ritual behaviours. We understand the ritual hand-cleanser not only because we find keeping clean important, but because we understand that expressing the importance of such a thing in our lives can be as important (in a different way) as the cleanliness itself.

Indeed, even if hand-clasping were explained simply as a tradition, without further explanation, we could understand its importance. Many people take the upholding of tradition to be an end in itself, quite independently of any thoughts about its benefits. In fact, people are prepared to sacrifice much in order to uphold tradition.

As Wittgenstein points out, Frazer is, in all likelihood, right that such rites must have had some initial significance in the lives of those who began the tradition, in order for the tradition to 'catch on' but we can imagine that those who engage in the practice now have no knowledge of the reasons for its inauguration (RFGB, pp. 148f). For some this

145 Of course, the example is not entirely appropriate, given that Christian baptism at least does, indeed, involve metaphysical beliefs about God and the afterlife, but one can also imagine a similar ritual that does not involve these beliefs.

146 See Schroeder, S. 2004 for a defence of the view that metaphors involve implicit comparisons.

may make the tradition seem 'pointless', and, indeed, most traditions are 'pointless', inasmuch as they serve no end outside themselves. For some, however, traditions are worth preserving for their own sake. This is an understandable human reaction. We all of us have traditions, rituals, routines, which are not undertaken for any further reason – they are just 'how we do things'.

An understanding of the attitudes of others doesn't require us to share the self-same attitude as our moral opponent in order to gain insight into it. I even share some common ground with those whose moral attitudes I vehemently disagree with. I may not, for example, think homosexual activities immoral, but such activities are not a matter of complete indifference to me in the way that the clasping and unclasping of hands is. Our sexual preferences are of fundamental importance in our lives. When we consider our own preferences and instinctive reactions, we can come to understand how these might manifest themselves in others as moral belief, even where we vehemently disagree with such a moral attitude.

Foot's example, then, is *prima facie* convincing against Humean liberalism not because the response in the case of the hand-clasper is not based on a particular kind of reason, for when we look, we can find examples of perfectly intelligible attitudes which take no account of such reasons. What accounts for the difference between such cases and Foot's hand-clasper, however, is that the response of the hand-clasper is one which we struggle to understand as a normal human response at all. As the example is set up, there is simply no connection whatsoever to be found between the hand-clasper's ritual and any instinct which we share.

Allan Gibbard claims that when we think of Foot's hand-clasper –:

> We're perhaps like the Masai boy in Kenya who flagged down a car with William Frankena in it with fellow bird-watchers: “What are they doing?” he

> asked the driver. “Looking for birds.” – “Oh, to eat them?” – “No.” Then after a pause, “Oh, they want the feathers” – “No.”) Like the Masai boy, we're baffled with the hand-clasper, because he's in a state of mind it's hard to imagine “from the inside”, even in mental play-acting. (Gibbard 2003, p. 28)

I want to say, however, that our bafflement with the hand-clasper is not like the Masai boy's bafflement at the bird-watchers (or, at least, it very likely isn't.) The Masai boy might very well not share the bird-watchers' fascination with birds, himself being interested in them for purely instrumental reasons, but it is unlikely that he is so far removed from the bird-watchers that, even on reflection, he would remain baffled by such a fascination. After all, it is a common human trait to be captivated by the natural world, if not by birds then, say, by mountains or stars. Hence the Masai boy, if he reflects, will be able to see a kinship between himself and the bird-watchers, however distant.

Aesthetic Appreciation and Moral Judgement

But moral judgements needn't only be connected with very basic, animal instincts, like cleanliness or sexuality in order to be understandable as such.

Wittgenstein often stressed the connection between ethics and aesthetics, famously claiming in the *Tractatus* that

> Ethics and aesthetics are one and the same. (TLP 6.422)

Indeed, the connection in Wittgenstein's mind between ethics and aesthetics can be seen even in his sole comment on ethics in the *Investigations*, where he talks about “our

concepts in ethics or aesthetics" (PI §77).

It is not hard to find cases where the line between the moral and the aesthetic is less than sharp. Manners, etiquette, and codes of dress are everyday examples. Of course, it is possible for such practices to have various grounds. Some may adhere to such social mores on the grounds that it is useful to do so, adapting their habits to different cultures and social settings. Others may do so out of respect for those who adhere to these rules on other grounds. But the attitude towards such practices that I am concerned with might best be summed up via the notion of 'seemliness'. An English missionary, for example, might try to inculcate a tribe with western social graces simply because it is 'unseemly' to act otherwise.

If we imagine that the hand-clasper's attitude towards hand-clasping is similar to the missionary's attitude towards, say eating with a knife and fork, its moral aspect becomes immediately understandable, even though hand-clasping is not a practice which appeals to us in the slightest.

This, however, raises once again the question of what marks out such an attitude as a moral one, as opposed, simply to a strange compulsion or affectation. For it is difficult to find here any particular connection to a shared human instinct. Of course, that there should be rules of some kind around how one eats is understandable, but there seems no explanation for why one should settle upon just these particular rules, and hence why adherence to the rules is any more than a preference.

Wittgenstein was similarly concerned with how a genuine aesthetic judgement was distinguished from a mere matter of taste. Whilst, according to Wittgenstein, expressions of aesthetic appreciation are related to simple interjections of approbation and disapprobation, usually first taught us as substitutes for our natural responses, say, to food (LC, p. 2), a genuine response of aesthetic appreciation is not such a simple interjection.

In his lectures on aesthetics, Wittgenstein says:

> The words we call expressions of aesthetic judgement play a very complicated role, but a very definite role, in what we call a culture of a period. To describe their use or to describe what you mean by a cultured taste, you have to describe a culture. (LC, p. 6)

According to Wittgenstein, genuine aesthetic judgements, as opposed to simple judgements of taste, require one's participation in a culture which has a widely shared system of rules according to which artworks are praised or criticised.

This does not mean, however, that aesthetic judgements are defined in terms of their objects or contents, or the particular kinds of reason which we take to justify them. African art may involve quite different rules or standards of judgement than does European art, but the judgements of an African artist or critic are still considered aesthetic ones (even if we do not appreciate his art), not because of the particular rules he follows, but because of the kind of place these rules play in the culture of which he is a member (LC, pp. 8f).

This isn't to say that one's judgements must always conform to rules shared within one's culture. In order to be considered competent, a composer must follow certain rules of composition to a great extent, but, as Wittgenstein acknowledges, “every composer changed the rules” (LC, p. 6).

Moral judgements, too, are intelligible only against a certain cultural background. One important aspect of the place of moral attitudes in a culture is that they are generally widely shared and passed down from one generation to the next. Thus the extent to which one's morals can be entirely individualistic is to some extent constrained.

This isn't, however, to say that in order to make a genuinely moral judgement one must always judge in conformity with others, or according to rules which one has been taught, as argued, from a Wittgensteinian perspective, by R.W. Beardsmore (1969, pp.

132ff). Beardsmore argues that moral decisions can only be made by reference to principles that one has been trained to accept, such that a moral decision only makes sense where the rules one has been taught come into conflict with one another. Hence, Beardsmore thinks, a 'decision of principle' makes no sense because one could not proffer any reason for the decision.

But this must also be true of decisions between conflicting rules, where we have no rule according to which the conflict must be settled one way or the other. Hence, if a decision to adopt a rule is unintelligible, then a decision to choose between the dictates of two conflicting rules is equally so.

The possibility of adopting new rules is conditioned more subtly than Beardsmore suggests. It is not that one can only have a moral response to actions which fall under certain rules in accordance with which one has antecedently been trained to respond. Rather, in order for one's aberrant responses to be taken seriously as genuine moral judgements, one must first be considered to have an appreciation of the mores of one's social group. That is, it is only the aberrant responses of those who, by and large, participate in a moral culture that are taken seriously enough to call 'moral'. Otherwise, they will be called mere 'preferences', 'tastes', or even, perhaps 'pathologies'.

None of this, however, is to say that moral judgements cannot be explained subjectively as expressions of attitude. Diverse standards are still allowed for, and any particular judgement might be rejected or revised. Rather, it is to say that Foot is correct to think that there are constraints on what might legitimately be called a moral attitude. The hand-clasper might behave in what, under one aspect, could be described as a precisely similar way towards hand-clasping outside of the context of involvement in a moral community, yet we would not call the response a moral one.

The truth in Foot's account, then, is that, in order for a response to be called a moral

judgement, a background is needed. But there is no neat way to define this background. In order for a response to be a moral one, rather than just a strong (and possibly pathological) preference, it must inhabit a particular place within a culture, which may be quite impossible to describe in all its detail (LC, p. 7), but which, as users of moral concepts, we nonetheless recognize.

A reading of Wittgenstein's *Remarks on Frazer's Golden Bough*, then, with its ethical implications in mind, can help to alleviate the puzzlement that many feel regarding the anti-realist's liberalism. The understanding of diverse moral viewpoints sought by opponents of liberalism can be found by recognising various 'kinships' between our own responses and those of others, and where these are lacking, we can nonetheless understand a response as a moral one when it occupies a certain characteristic place within a moral culture. Thus Foot is to some extent vindicated. A moral response does indeed require a background, although not the background which Foot claimed.

Conclusion

Wittgenstein once said that he wanted his philosophy to "get something done" (Drury 1984, p. 110). In this thesis, I have tried to outline a metaethic, which, I hope, does just that – a metaethic, that is, which is useful.

Confusion around the use of moral concepts is perhaps more apt than most cases of philosophical error to lead us astray in our lives, and that in a most grave way. For decisions with a moral aspect have a special kind of gravity. Wrong-headed instrumental reasoning may well make one's life more difficult in many ways, but wrong-headed moral thinking makes it difficult to live with oneself.

As I have argued, Wittgenstein's philosophy can help us to avert moral error, as it can help us to avert error in other areas, such as the sciences and mathematics, by ensuring that our beliefs and attitudes are not based on confused philosophical presuppositions.

In attempting to find a path away from philosophical error in moral thinking, I have taken Wittgenstein as a guidepost. His thoughts on ethics point in the direction of a broadly expressivist treatment of moral terms, and also serve to highlight some of the confusions that such a metaethic can itself bring upon us: that there is a conceptual connection between moral judgements and conative or affective responses, yet that moral status is not ultimately contingent upon such responses, that the form of our language makes moral judgements look so much like factual ones and that we are confused as to how moral disagreement is possible if such judgements are conceived, instead, as expressions of attitude.

Given Wittgenstein's relative silence on ethics, though, his philosophy can *only* be a guidepost. Thus we need to look elsewhere than his sparse contributions to ethics if we are to make out the trail. Wittgenstein's comments on psychological concepts help us to do

this, by dispelling the notion that an expressive account of ethics reduces it to a kind of shouting match of 'Boos!' and 'Hurrahs!', and also provides us with a model for ethical pronouncements in showing us how an expressive utterance can *also* be descriptive.

Whilst the notion that moral judgements are expressions of attitude is an important insight, though, some philosophers, in attempting to follow Wittgenstein, have perhaps gone too far down this path and into dangerous territory. For the non-cognitivist turn has not yet led us to our destination, and in the meantime it has ushered us into confusions as thorny as those which it was meant to help us escape.

With the light of Wittgenstein's reminder that meaning is use, however, we may yet be able to find our way back onto a clearer path. For this shows that the apparent dangers which the non-cognitivist left the road in order to avoid were, in fact, chimeras. With their 'parity thesis' expressivists rightly (if dimly) recognised that, whilst '*x* is wrong' and 'I disapprove of *x*' have the same simple assertive use, they nonetheless do not share the same overall use, much as '*p*' and 'I believe that *p*'. The same, however, can be said of 'I disapprove of *x*' and '*x* is such that I would disapprove of it after careful consideration'. Both are expressions of disapproval when used atomically, but they cannot be used interchangeably in embedded contexts.

The non-cognitivist may be right to point out that '*x* is wrong' and '*x* is such that I would disapprove of it after careful consideration' cannot be used interchangeably in every construction either. For whilst 'If *x* weren't wrong, then *x* wouldn't be wrong' is a bald tautology, 'If *x* weren't such that after careful consideration I would disapprove of it, then *x* wouldn't be wrong' is, at least on one reading, false.

To reject the analysis on the grounds of such a minimal difference in the uses of these sentences, however, would be to throw the baby out with the bath water. Whilst the non-cognitivist sees his observation as a fatal flaw in the subjectivist analysis, with a

correct view of the aim of a philosophical analysis, we can simply add the rider that the uses of these sentences differ in this way, and retain the usefulness of the analysis in other respects. Thus the 'modal problem' should not be viewed as an objection to a philosophical theory, but as a benign grammatical observation, which serves to forestall our abandoning an otherwise enlightening synopsis of the grammar of moral terms.

Subjectivism's second nemesis – the disagreement problem – on the other hand, is shown to be the result of a simple misunderstanding regarding the use of the truth-predicate. That an utterance made in terms of our analysis may differ in truth-value from an utterance couched in moral terms serves not to show a difference in meaning between *analysandum* and *analysans*, but merely to highlight a fact about the meaning of 'true' and its cognates, namely that they are not used to say something *about* a sentence, but, in essence, merely to *repeat* it, or to utter its negation, transposed into reported speech.

Perhaps, though, we lack a rationale for taking our convoluted path. Perhaps our twists and turns are taken in the attempt merely to get us to where we were determined to go, rather than towards the goal we should have been seeking – the truth. And perhaps, even if the philosopher is justified in taking this path, there is no rationale for following him in using a concept thus described.

These qualms, too, though, are born of confusion. For we did not adopt our analysis only in order to rig a philosophical game, but in order correctly to describe our practice with moral terms, and there can be no other criterion of correctness for a philosophical analysis. That an analysis has been altered to fit the linguistic facts, then, can be no objection to it. Moreover, as moral agents, we are concerned to act in ways that can withstand rational scrutiny. We should be careful what we wish for because things may not be as they appear, and we may end up getting something other than we bargained for. Thus there is every reason to use moral concepts as our analysis portrays them.

There are those, however, who argue that our map of moral concepts is at best incomplete, and at worst fails to represent them at all, allowing such absurdities as a moral penchant for the clasping and unclasping of one's hands at regular intervals. Wittgenstein's reflections on ritual practices and aesthetics, however, show that we are not committed to such a recklessly liberal picture. A moral judgement, in order to be intelligible to us as such, must bear *some* relation to attitudes that we too share, or else, at least be issued by someone who is a participant in a recognizable moral culture which shares kinships with our own.

Idealized subjectivism, then, needn't make ethics objectionably 'mind-dependent' or absurdly 'liberal'. Neither need it be an objectionably *ad hoc* fix for an ailing theory. Instead, it is simply a reminder of what, as moral agents, we know implicitly – that when making decisions about how to live, we should be careful what we wish for, as failing to do so can lead us to act in ways which we may later regret. In arguing thus, I hope, in the end, to have made an old metaethic useful.

Bibliography

Allen, C. & Trestman, M. (2015). 'Animal Consciousness'. In Zalta, E. (ed.). *The Stanford Encyclopedia of Philosophy*. [WWW document] retrieved 6th January 2016 from the World Wide Web: http://plato.stanford.edu/cgi-bin/encyclopedia/archinfo.cgi?entry=consciousness-animal

Arrington, R. (1989). *Rationalism, Realism and Relativism*. New York: Cornell University Press.

Ayer, A. (1936). *Language, Truth and Logic*. London: Penguin.

Baker, G. (2004). *Wittgenstein's Method: Neglected Aspects*. Oxford: Blackwell.

Baker, G. & Hacker, P. (1980a). *Wittgenstein: Understanding and Meaning – Volume 1 of an Analytical Commentary on the Philosophical Investigations. Part I – Essays*. 2nd edn. (2009) Malden, Mass.; Oxford; Chichester: Wiley-Blackwell.

Baker, G. & Hacker, P. (1980a). *Wittgenstein: Understanding and Meaning – Volume 1 of an Analytical Commentary on the Philosophical Investigations. Part II – Exegesis §§1-184*. 2nd edn. (2009) Malden, Mass.; Oxford; Chichester: Wiley-Blackwell.

Baker, G. & Hacker, P. (1985). *Wittgenstein: Rules, Grammar and Necessity – Volume 2 of An Analytical Commentary on the Philosophical Investigations. Essays and Exegesis of §§185-242.* 2nd edn. (2009) Malden, Mass.; Oxford; Chichester: Wiley-Blackwell.

Beardsmore, R. (1969). *Moral Reasoning: Vol II*. London; New York: Routledge.

Bennet, M. & Hacker, P. (2003). *The Philosophical Foundations of Neuroscience.* Malden; Oxford; Carlton: Blackwell.

Black, M. (1964). *A Companion to Wittgenstein's Tractatus*. Cambridge: Cambridge University Press.

Blackburn, S. (1981). 'Reply: Rule-Following and Moral Realism'. In Holtzman, S. & Leich, C. (eds.). *Wittgenstein: To Follow a Rule*. London: Routledge.

Blackburn, S. (1984). *Spreading the Word*. Oxford: Oxford University Press.

Blackburn, S. (1990). 'Wittgenstein's Irrealism'. Reprinted in his (2010). *Practical Tortoise Raising: and other philosophical essays*. Oxford: Oxford University Press.

Blackburn, S. (1993). *Essays in Quasi-Realism*. New York: Oxford University Press.

Blackburn, S. (1998). 'Wittgenstein, Wright, Rorty and Minimalism'. In *Mind*, New Series. 107(425):157-181.

Boisvert, D. (2015). 'Charles Leslie Stevenson'. In Zalta, E. (ed.). *The Stanford Encyclopedia of Philosophy*. [WWW document] retrieved 6th January 2016 from the World Wide Web: http://plato.stanford.edu/entries/stevenson/

Borg, E. (2010). 'Meaning and Context: A Survey of a Contemporary Debate'. In Whiting, D. (ed.). *The Later Wittgenstein on Language*. Basingstoke; New York: Palgrave Macmillan. pp. 96-113.

Bouwsma, O. (1961). 'The Blue Book'. In *The Journal of Philosophy.* 58(6):141-162.

Cavell, S. (1999) T*he Claim of Reason: Wittgenstein, Skepticism, Morality and Tragedy*. New York: Oxford University Press.

Christensen, A. (2011). 'Wittgenstein and Ethics'. In Kuusela, O. & McGinn, M. (eds.). *The Oxford Handbook of Wittgenstein*. [WWW document] retrieved 6th January 2016 from the World Wide Web: http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199287505.001.0001/oxfordhb-9780199287505-e-36

Cottingham, J. (1978). 'A Brute to the Brutes?: Descartes' Treatment of Animals'. In *Philosophy.* 53(206):551-559.

Dancy, J. (1993). *Moral Reasons*.Oxford; Cambridge, Mass.: Blackwell.

Dancy, J. (2004). *Ethics Without Principles*. Oxford; New York: Oxford University Press.

Dancy, J. (2013). 'Moral Particularism'. In Zalta, E. (ed.). *The Stanford Encyclopedia of Philosophy.* [WWW document] retrieved 23rd October 2016 from the World Wide Web: http://plato.stanford.edu/entries/moral-particularism/

Delapp, K. (2016). 'Metaethics'. In Fieser, J. & Dowden, B. (eds.) *Internet Encyclopedia of Philosophy.* [WWW document] retrieved 23rd October 2016 from the World Wide Web: http://www.iep.utm.edu/metaethi/#SH1b

Diamond, C. (1981). 'What Nonsense Might Be'. In *Philosophy* 56(215):5-22.

Dreier, J. (2004). 'Lockean and Logical Truth Conditions.' In *Analysis* 64(1):84-91.

Drury, M. (1984). 'Conversations With Wittgenstein'. In Rhees, R. (ed.) *Recollections of Wittgenstein.* Oxford; New York: Oxford University Press. pp. 97-171.

Edmonds, D. & Eidinow, J. (2002). *Wittgenstein's Poker.* New York: Harper Collins.

Enoch, D. (2005). 'Why Idealize?'. In *Ethics.* 155(4): 759-787.

Enoch, D. (Unpublished). 'Idealizing Still Not Off the Hook: A Reply to Sobel's Reply'. [WWW document] Downloaded 20th August 2016 from the World Wide Web: http://law.huji.ac.il/upload/ReplyToSobel.doc

Foot, P. (1958). 'Moral Beliefs'. In *Proceedings of the Aristotelian Society.* 59:83-104.

Frazer, J. (1890) *The Golden Bough*. (New Abridgement) (1998). Oxford: Oxford University Press.

Geach, P. (1958). 'Imperative and Deontic Logic'. In *Analysis* 18(3): 49-56.

Geach, P. (1965). 'Assertion'. In *The Philosophical Review* 74(4):449-465.

Gibbard, A. (1990). *Wise Choices, Apt Feelings*. Oxford: Oxford University Press.

Gibbard, A. (2003). *Thinking How to Live*. Cambridge: Harvard University Press.

Glock, H. (1991). 'Philosophical Investigations Section 128: “Theses in Philosophy” and Undogmatic Procedure'. In Arrington, R. & Glock, H. (eds.). *Wittgenstein's Philosophical Investigations: Text and Context*. London; New York: Routledge.

Glock, H. (1996). 'Abusing Use'. In *Dialectica* 50(3):205-224.

Glock, H. (2000). 'Animals, Thoughts and Concepts'. In *Synthese* 123(1):35-64.

Glock, H. (2004). 'All Kinds of Nonsense'. In Ammereller, E. & Fischer, E. (eds.). *Wittgenstein at Work: Method in the Philosophical Investigations*. London; New York: Routledge. pp. 221-245.

Glock, H. (2004). 'Knowledge, Certainty and Scepticism: In Moore's Defence'. In Moyall-Sharrock, D. (ed.) *The Third Wittgenstein: The Post-Investigations Works*. Aldershot: Ashgate.

Glock, H. (2015). 'Wittgensteinian Anti-Anti Realism: One 'Anti' Too Many?'. In *Ethical Perspectives*. 22(1):99-129.

Grayling, A. (1988). *Wittgenstein*. Oxford: Oxford University Press.

Grice, P. (1989). *Studies in the Way of Words*. Cambridge Mass.: Harvard University Press.

Hacker, P. (1987). *Appearance and Reality: A Philosophical Inquiry into Perception and Perceptual Qualities*. Oxford; New York: Basil Blackwell.

Hacker, P. (1990a). *Wittgenstein: Meaning and Mind – Volume 3 of an Analytical Commentary on the Philosophical Investigations. Part I – Essays*. Oxford; Malden (Mass.): Blackwell.

Hacker, P. (1990b). *Wittgenstein: Meaning and Mind – Volume 3 of an Analytical Commentary on the Philosophical Investigations*. Oxford; Malden (Mass.): Blackwell.

Hacker, P. (2008). 'Gordon Baker's Late Interpretation of Wittgenstein'. In Kahane, G. et al. (eds.) *Wittgenstein and His Interpreters: Essays in Memory of Gordon Baker.* Ames, Iowa: Blackwell Publishing Ltd. pp. 88-122.

Hare, R. (1952). *The Language of Morals*. London: Oxford University Press.

Hare, R. (1965). *Freedom and Reason*. London: Oxford University Press.

Heal, J. (1994). 'Moore's Paradox: A Wittgensteinian Approach'. In *Mind*. 103(409):5-24.

Horwich, P. (1993). 'Gibbard's Theory of Norms'. In *Philosophy and Public Affairs*. 22:67-78.

Hume, D. (1739). *A Treatise of Human Nature*. (1969). Harmondsworth; Baltimore; Ringworth: Penguin Books.

Jackson, F. & Pettit, P. (1998). 'A Problem for Expressivists'. In *Analysis*. 58:239-251.

Johnson, P. (1989). *Wittgenstein and Moral Philosophy.* London; New York: Routledge.

Johnston, M. (1989). 'Dispositional theories of value'. In *Proceedings of the Aristotelian Society, Supplementary Volumes*. 63:139-174.

Kenny, A. (1986). 'Wittgenstein on the Nature of Philosophy'. Reprinted in Shanker, S. & Kilfoyle, D. (eds.) (2002) *Ludwig Wittgenstein: Critical Assessments of Leading Philosophers. Vol. II – The Later Wittgenstein: From Philosophical Investigations to On Certainty*. London; New York: Routledge. pp. 7-25.

Lewis, D. (1989). 'Dispositional theories of value'. In *Proceedings of the Aristotelian Society, Supplementary Volumes*. 63:113-137.

Lewy, D. (2007). 'Wittgenstein's Philosophical Work on Ethics'. In Zamuner, E., Di Lascio, V. & Lewy, D. (eds.). *Ludwig Wittgenstein: Lecture on Ethics, Introduction, Interpretation and Complete Text.* Macerata: Quodlibet Società Cooperitiva.

Lovibond, S. (1983). *Realism and Imagination in Ethics*. Oxford: Basil Blackwell.

Lycan, W. (2000). *Philosophy of Language*. London: Routledge.

Macarthur, D. (2010). 'Wittgenstein and Expressivism'. In Whiting, D. (ed.). *The Later Wittgenstein on Language*. Basingstoke; New York: Palgrave Macmillan. pp. 81-95.

Mackie, J. (1977). *Ethics: Inventing Right and Wrong*. Harmonsdworth: Penguin.

Magee, B. (1983). *The Philosophy of Schopenhauer*, Oxford: Clarendon Press.

Malcolm, N. (1995). 'Disentangling Moore's Paradox'. In Egidi, R. (ed.). *Wittgenstein: Mind and Language*. Synthese Library (Vol. 245). pp. 195-205.

Marcuse, H. (1964). *One Dimensional Man*. Abingdon; New York: Routledge.

McDowell, J. (1979). 'Virtue and Reason'. In *The Monist*. 61:331-350.

McDowell, J. (1985). 'Values and Secondary Qualities'. In Honderich, T. (ed.). *Morality and Objectvity.* Reprinted in McDowell, J. (1998). *Mind, Value and Reality*. Cambridge, Mass; London: Harvard University Press.

McDowell, J. (1990). 'Non-cognitivism and Rule-Following'. In Holtzman, S. & Leich, C. (eds.). *Wittgenstein: To Follow a Rule*. London: Routledge.

McGuinness, B. (1988). *Wittgenstein, A Life: Vol.1 Young Ludwig 1889-1921*. London: Duckworth.

Mill, J. (1859). *On Liberty.* (1982) London: Penguin Books.

Monk, R. (1990). *Ludwig Wittgenstein: The Duty of Genius.* London: Cape.

Moore, G. (1903). *Principia Ethica*. Cambridge: Cambridge University Press.

Moore, G. (1912). *Ethics*. 2nd edn. (1966). London: Oxford University Press.

Moore, G. (1942). 'A Reply to My Critics'. In Schilpp, p. (ed.). *The Library of Living Philosophers (Vol IV): The Philosophy of G.E. Moore*. Evanston; Chacago: Northwestern University. pp. 535-677.

Moore, G. (1954). 'Wittgenstein's Lectures in 1930-33'. In *Mind, New Series*. 63 (249):1-15.

Moore, G. (1955). 'Wittgenstein's Lectures in 1930-33'. In *Mind, New Series*. 64 (253):1-27.

Moore, G. (1959). *Philosophical Papers*. London: George Unwin.

Ogden, C. & Richards, I. (1923). *The Meaning of Meaning*. New York: Harcourt.

Phillips, D. & Mounce, H. (1969). *Moral Practices*. London: Routledge.

Pleasants, N. (2009). 'Wittgenstein and Basic Moral Certainty'. In *Philosophia*. 37:669-679.

Popper, K. (1952). 'The Nature of Philosophical Problems and their Roots'. In *The British Journal for the Philosophy of Science* 3(10):124-156.

Ramsey, F. (1927). 'Facts and Propositions'. In Mellor, D. (ed.). *Philosophical Papers*. Cambridge: Cambridge University Press. pp. 34-51.

Rhees, R. (1965). 'Some Developments in Wittgenstein's View of Ethics'. In *The Philosophical Review.* 74:17-26.

Rhees, R. (ed.) (1984). *Recollections of Wittgenstein.* Oxford; New York: Oxford University Press.

Ridge, M. (2006). 'Sincerity and Expressivism'. In *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*. 131(2):487-510.

Ross, W.D. (1939). *Foundations of Ethics*. London: Oxford University Press.

Rundle, B. (1990). *Wittgenstein and Contemporary Philosophy of Language*. Oxford: Blackwell.

Russell, B. (1959). *My Philosophical Development*. London: Unwin Books.

Schopenhauer, A. (1818). *The World as Will and Idea*. Haldane, R. & Kemp, J. (trans.) (1909). London: Kegan-Paul.

Schroeder, M. (2008). *Being For.* Oxford; New York: Oxford University Press.

Schroeder, M. (2009). 'Hybrid Expressivism: Virtues and Vices'. In *Ethics* 119:257-309.

Schroeder, M. (2010a). *Noncognitivism in Ethics*. Abingdon; New York: Routledge.

Schroeder, M. (2010b). 'Does Expressivism Have Subjectivist Consequences?'. [WWW Document]. Retireved 16 Arpil 2013 from the World Wide Web: http://www-bcf.usc.edu/~maschroe/research/Schroeder_Subjectivist_Consequences.pdf

Schroeder, M. (2014). 'Does Expressivism Have Subjectivist Consequences?'. In *Philosophical Perspectives*. 28:278-90.

Schroeder, S. (2004). 'Why Juliet is the Sun'. In Siebel, M & Textor, M. (eds.). *Semantik und Ontologie*. Frankfurt; Main: Ontos Verlag. pp. 63-101.

Schroeder, S. (2006a). *Wittgenstein: The Way Out of the Fly-Bottle*. Cambridge; Malden, Mass.: Polity Press.

Schroeder, S. (2006b). 'Moore's Paradox and First Person Authority'. In *Grazer Philosophische Studien*. 71:161-174.

Schroeder, S. (2007). 'The Tightrope Walker'. In *Ratio (New Series).* 20(4):442-463.

Schroeder, S. (2009). 'Analytic Truths and Grammatical Propositions'. In Glock, H. & Hyman, J. (eds.). *Wittgenstein and Analytic Philosophy: Essays for P.M.S. Hacker.* Oxford; New York: Oxford University Press.

Schroeder, S. (2012). 'Schopenhauer's Influence on Wittgenstein'. In Vandenabeele, B. (ed.), *A Companion to Schopenhauer.* Oxford: Wiley-Blackwell. pp. 367-384.

Schroeder, S. (Forthcoming). 'Wittgenstein on Aesthetics'. In Glock, H. & Hyman, J. (eds.). *A Companion to Wittgenstein*. Blackwell.

Searle, J. (1969). *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.

Searle, J. (1978). 'Literal Meaning'. In *Erkenntnis.* 13(1):207-224.

Searle, J. (2007). 'Grice on Meaning: 50 Years Later'. In *Teorema* 26(2):9-18.

Shafer-Landau, R. (2003). *Moral Realism: A Defence*. Oxford: Oxford University Press.

Sinclair, N. (2009). 'Recent Work in Expressivism'. In *Analysis Reviews*. 69(1):136-147.

Sinclair, N. (2012). 'Moral Realism, Face-Values and Presumptions'. In *Analytic Philosophy*. 53(2):158-179.

Smith, M. (1986). 'Should We Believe in Emotivism?'. In McDonald, G. & Wroght, C. (eds.). *Fact, Science and Morality: Essays on A.J. Ayer's Language, Truth and Logic*. Oxford: Basil Blackwell.

Sobel, D. (2009). 'Subjectivism and Idealization'. In *Ethics*. 119(2):336-352.

Stern, D. (2013). 'Wittgenstein's Lectures on Ethics, Cambridge 1933'. In *Wittgenstein-Studien*. 2(1):191-205.

Stevenson, C. (1937). 'The Emotive Meaning of Ethical Terms'. In *Mind*. (New Series). 46(181):14-31.

Stevenson, C. (1942). 'Moore's Arguments against Certain Forms of Ethical Naturalism'. In Schilpp, P. (ed.). *The Library of Living Philosophers (Vol IV): The Philosophy of G.E. Moore*. Evanston; Chacago: Northwestern University. pp. 69-90.

Stevenson, C. (1944) *Ethics and Language*. Clinton: Yale University Press.

Suikkanen, J. (2009). 'The Subjectivist Consequences of Expressivism.' In *Pacific Philosophical Quarterly*. 90(3):364-387.

Szabados, B. (1999). 'Was Wittgenstein an Anti-Semite?: The Significance of Anti-Semitism for Wittgenstein's Philosophy'. In *Canadian Journal of Philosophy*. 29(1):1-28.

Szabados, B. (2010). *Ludwig Wittgenstein on Race, Gender and Cultural Identity*. New York: Edwin Mellen.

Unwin, N. (1999). 'Quasi-realism, Negation and the Frege-Geach Problem'. In *The Philosophical Quarterly*. 49(196):337-352.

Unwin, N. (2001). 'Norms and Negation: A Problem for Gibbard's Logic'. In *the Philosophical Quarterly*. 51(202):60-75.

Weiner, D. (1992). *Genius and Talent: Schopenhauer's Influence on Wittgenstein's Early Philosophy*, Rutherford: Fairleigh Dickinson University Press.

Wiggins, D. (1987). *Needs, Values, Truth: Essays in the Philosophy of Value*. 3rd end. (2002). Oxford; New York: Oxford University Press.

von Wright, G. (1982). 'The Wittgenstein Papers'. In his *Wittgenstein*. Oxford: Blackwell. pp. 35-62.

Texts by Ludwig Wittgenstein (Abbreviations)

Wittgenstein, L.

(1979). *Wittgenstein's Lectures: Cambridge 1932-1935*. Oxford: Blackwell. (**AWL**)

(1969). *The Blue and Brown Books: Preliminary Studies for the Philosophical Investigations*. (2nd edn.). Oxford: Blackwell. (**BB**)

(2005). *The Big Typescript*. Oxford: Blackwell. (**BT**)

(1995). *Cambridge Letters: Correspondence with Russell, Keynes, Moore, Ramsey and Sraffa*. Oxford: Blackwell. (**CL**)

(1998). *Culture and Value*. (Revised. edn.). G. H. von Wright (ed.). P. Winch (Trans.). Oxford: Blackwell. (**CV**)

(1984). *Eine Philosophische Betrachtung*. In *Werkausgabe Band, 5*. Frankfurt; Main: Shurkamp. pp. 117-282. (**EPB**)

(1992). *Geheime Tagebücher, 1914-1916.* 3rd edn. Baum, W. (ed.). Vienna: Turia & Kant. (**GT**)

(1970). *Lectures and Conversations on Aesthetics, Psychology and Religious Belief*. Oxford: Basil Blackwell. (**LC**)

(1965). 'A Lecture on Ethics'. In *The Philosophical Review* 74(1): 3-12. (**LE**)

(1979). 'Letters to Ludwig von Ficker'. In Luckhardt, C. (ed.). *Wittgenstein: Sources and Perspectives.* Hassocks: Harvester Press. (**LF**)

(1988). *Lectures on Philosophical Psychology – 1946-7: Notes by P.T. Geach, K.J. Shah, A.C. Jackson*. Geach, P. (ed.). Hemel Hempstead: Harvester Wheatsheaf. (**LPP**)

(1984). 'The language of Sense Data and Private Experience'. In *Philosophical Investigations*. 7(1):1-45. (**LSD**)

(1982). *Last Writings on the Philosophy of Psychology (Vol I)*. Oxford: Blackwell. (**LW I**)

(19p2). *Last Writings on the Philosophy of Psychology (Vol II)*. Oxford: Blackwell. (**LW II**)

Unpublished manuscripts, numbered in accordance with von Wright (1982). (**MS**)

(1998). *Notebooks 1914-16*. Malden, MA; Oxford; Carlton: Blackwell. (**NB**)

(1969). *On Certainty*. New York; Hagerstown, San Francisco; London: Harper & Row. (**OC**)

(1974). *Philosophical Grammar*. Oxford: Blackwell. (**PG**)

(1978). *Philosophical Investigations*. Oxford: Basil Blackwell. (**PI**)

(1975). *Philosophical Remarks*. Oxford: Blackwell. (**PR**)

(1993). 'Remarks on Frazer's Golden Bough'. In Klagge, J. & Nordmann, A. (eds.). *Philosophical Occasions 1912-1951*. Indianapolis: Hackett. pp. 119-55. (**RFGB**)

(1956). *Remarks on the Foundations of Mathematics*. Oxford: Basil Blackwell. (**RFM**)

(1929). 'Some Remarks on Logical Form'. In *Proceedings of the Aristotelian Society, Supplementary Volumes*. 9:162-71. (**RLF**)

(1980). *Remarks on the Philosophy of Psychology (Vol 1)*. Oxford: Basil Blackwell. (**RPP I**)

(1980). *Remarks on the Philosophy of Psychology (Vol II)*. Oxford: Basil Blackwell. (**RPP II**)

(2001). *Tractatus Logico-Philosophicus*. London; New York: Routledge. (**TLP**)

Unpublished typescripts, numbered in accordance with von Wright (1982). (**TS**)

(1967). *Zettel*. Oxford: Basil Blackwell. (**Z**)