**EXPRESSIVISM: A CONCISE DEFENCE**

**By**

**HOWARD SIMMONS**

**TABLE OF CONTENTS**

1. Introduction
	1. Expressivism 4
	2. Mixed normative statements 8
	3. The Redundancy Thesis 12
2. Expressivism: The Basics
	1. Expressing attitudes 15
	2. Geach on ‘good’ 17
	3. Deontic words 20
	4. Speech and thought 22
3. Expressivism: Truth and the Embedding Problem

3.1 Geach’s challenge 24

3.2 Inferentialism 26

3.3 Propositional attitudes 33

3.4 Is normative truth relative? 37

3.5 Parfit on truth in expressivism 39

1. Morality: structure and content

4.1 Moral statements 43

4.2 Structural connections 45

4.3 Substantive features 53

1. Alternative theories and their shortcomings

5.1 Introduction 56

5.2 Basic naturalism 57

5.3 ‘Cornell realism’ 63

5.4 Intuitionism 65

5.5 The error theory 67

5.6 Constructivism 68

5.7 Hybrid theories 70

1. Morality: Theory and Rhetoric

6.1 The place of theory in ethics 75

6.2 Objectivity and rhetoric 79 6.3 Seay on theory scepticism and moral dilemmas 82

 7. Implications of expressivism

 7.1 Normative disagreement 85

 7.2 God’s values 87

 7.3 Conclusion 90

 Glossary 92

 Bibliography 94

**CHAPTER ONE**

**Introduction**

**1.1 Expressivism**

This book is about normative statements.[[1]](#footnote-1) I claim that the distinguishing mark of a normative statement is that at least part of what it does is express an attitude on the part of the speaker, and that this feature may be considered an essential part of the meaning of the statement.[[2]](#footnote-2),[[3]](#footnote-3) The attitude may be positive (as in the case of ‘This wine is good’) or negative (as in the case of ‘That plan is bad’), or it may be comparative (as in the case of ‘A dog would be better for you than a cat’). Each of these cases admits of degrees (‘This is very good’, ‘That was somewhat bad’). The attitude may also be ‘presumptive’ or prima facie (as in ‘Since he told a lie, his action was in that way bad’) or may be held ‘decisively’ or ‘all things considered’ (as in ‘His action was bad overall’). A *normative term* is one whose use in a statement causes that statement to be normative, i.e., express an attitude. (Examples invoked so far include ‘good’, ‘bad’ and ‘better’.) This conception of what it is for a statement or term to be normative is called *expressivism*.

It is advisable to say something here about the relationship between attitudes and preferences. Attitude-talk and preference-talk are very closely related, though the latter is normally more circumscribed. I may have a preference for eating chocolate over not eating chocolate on a particular occasion, but if I have a positive *attitude* to chocolate, this is normally understood to involve more than just my eating it on a given occasion - it will also encompass my position on such things as others’ eating of chocolate and the merits of chocolate in general. However, we can ‘reduce’ preference-talk to attitude-talk by referring to ‘ideas’ or ‘prospects’. To have a preference, all things considered, for eating chocolate over not eating chocolate on some occasion is equivalent to having a more positive attitude, all things considered, towards the idea or prospect of (my) eating chocolate on that occasion than to the idea of (my) not eating chocolate on that occasion. Normative statements thus sometimes express preferences, as in ‘I should leave now’, which can be used to express a preference for the idea of my leaving over my not leaving at the relevant time. Throughout this book I shall mostly talk about attitudes, but when it seems more apt, I will sometimes talk of preferences, on the understanding that preferences are a species of attitude. As for attitudes themselves, I do not provide any formal definition of the term ‘attitude’. I will, however, draw attention to one point here. In general, an attitude has both a phenomenological component and a behavioural component. That is to say, it often *feels* like something to have a certain attitude, but it also has implications for the behaviour of the person whose attitude it is. For example, if someone has a negative attitude to cheating, then they will typically experience feelings of disappointment or contempt when considering acts of cheating by others and feelings of disappointment or shame if they engage in any cheating themselves. This is the phenomenological aspect of the attitude. But they will also be somewhat motivated not to *engage* in cheating, which is the behavioural aspect of the attitude.

Instead of positive and negative attitudes, I will sometimes speak of approving and disapproving of things, which I teat as synonymous.

Normative statements can be contrasted with descriptive statements, such as ‘Water is a compound’ or ‘Their lawn is in the back garden’, which do not express attitudes. This point is enough, I maintain, to distinguish between normative and descriptive statements. It would not be appropriate to add that descriptive statements represent or correspond to reality or have ‘mind-to-world fit’, since, as we shall see, this is true of some normative statements as well. Nor should we distinguish the two categories by saying that descriptive statements express beliefs, whereas normative statements express attitudes. For while it is true that normative statements express attitudes, whereas descriptive statements do not (by definition), there is also a sense in which normative statements express beliefs. The concept of a descriptive statement is framed here in a purely negative way: descriptive statements are the set of statements whose members do not express attitudes.

However, it might be argued that the distinction between normative and descriptive statements is fundamentally flawed, since descriptive statements, like normative ones, express attitudes. If I say, for example, ‘The oven is hot’, it might be claimed that I am in some way conveying the fact that I think that ‘hot’ is an appropriate word to describe the oven, and thereby expressing a positive attitude to the use of that word in this context. Even if, for some reason, I disliked the English language, and thought it should use completely different words, I am still implying in some sense that ‘hot’ is a good way to describe the oven according to the actual norms of the English language and the empirically determinable state of the oven. Is this sufficient reason for declaring that what we have been calling descriptive statements express attitudes after all, and thus for abandoning the distinction between descriptive and normative statements, at least in the way that I have framed it?

I do not believe it is. For the implied claim that the word ‘hot’ is a good word to describe the oven according to the norms of English is not a normative claim in the sense intended in this book. Though the speaker, in implying this claim, can be taken as saying that English has certain norms determining the use of its words and phrases, including the use of the word ‘hot’ in this situation, and that they themselves are *conforming* to those norms when they describe the oven in this way, they are not actually expressing any personal approval for this use of the word, as is shown by the fact that their conformity to the norms of English may be reluctant.[[4]](#footnote-4)

In any case, even if we did end up having to say that descriptive statements express attitudes to possible uses of words, this would not entail that we had to give up the distinction between normative and descriptive statements. For that distinction could simply be recast as the contrast between expressing attitudes concerning the *subject-matter* of the statement and expressing attitudes only about the use of the words within it when this is not part of the statement’s subject-matter.[[5]](#footnote-5)

We say that a simple, stand-alone normative statement *expresses* an attitude towards something, not that it asserts that the speaker has that attitude. The distinction between asserting and expressing is an important one for expressivists. It is a distinction which also applies to descriptive statements. “Water is a compound” asserts that water is a compound, but what it expresses is the *belief* that water is a compound. So having an attitude stands to making a normative statement as having a belief stands to making a descriptive statement. Without this distinction, the expressivist theory collapses into *speaker subjectivism*.[[6]](#footnote-6) For example, if Anya says ‘Capital punishment is wrong’, the theory will say that Anya is thereby asserting that she has a negative attitude to capital punishment, and that her statement is consequently logically equivalent to ‘I have a negative attitude to capital punishment’. This has several implausible consequences:

* According to speaker subjectivism, if Anya and Bill both say ‘Capital punishment is wrong’, then they are not asserting the same thing, for each is asserting of a *different* person that they have a negative attitude to capital punishment. (Anya is asserting that she does, and Bill is asserting that he does.)
* According to speaker subjectivism, if Anya says ‘Capital punishment is wrong’ and Bill says ‘Capital punishment is not wrong’, it appears that they do not really disagree, as Anya is saying that she has a negative attitude to capital punishment and Bill is saying that he doesn’t, and there is no logical inconsistency between these two claims (Schroeder 2008, 17).
* According to speaker subjectivism, the statement ‘If I didn’t have a negative attitude to capital punishment, then capital punishment wouldn’t be wrong’ must be a logical truth, since it is equivalent to ‘If I didn’t have a negative attitude to capital punishment, then I wouldn’t have a negative attitude to capital punishment’. Yet the former statement is surely *not* a logical truth (Schroeder 2008, 16).[[7]](#footnote-7)

Since expressivism does not claim that ‘Capital punishment is wrong’ is equivalent to ‘I have a negative attitude to capital punishment’, it is not subject to any of these problems. Notice that parallels to each of these problems would arise for someone who implausibly thought that ‘Water is a compound’ was equivalent to ‘I believe that water is a compound’. Seeing that descriptive statements express beliefs rather than asserting them would dissolve away these parallel problems, just as seeing that normative statements express attitudes dissolves away the original ones.

Although, as we have just seen, expressivism is different from speaker subjectivism, there is a sense in which an expressivist will treat all normative statements as ‘subjective’. For there is nothing which, so to speak, validates people’s subjective attitudes. To say that someone’s attitude is wrong is just to express another attitude. Note also that the expressivism advocated in this book is *global* in the sense that it apples to *all* normative statements. This includes moral statements and even statements about what factual statments we ought to believe.

But, apart from its superiority to speaker subjectivism, why do I accept the expressivist view of normative statements? The main reason is as follows. It is very plausible that (a) ‘That is a terrible joke’ and ‘Murder is wrong’, for example, share something which, say, ‘Water is a compound’ and ‘Their lawn is in the back garden’ lack; and that (b) this thing which they have in common is that they both express attitudes of some sort. It is true that many philosophers are sceptical about expressivism. One reason for this is that expressivism seems to treat normative statements as being more like exclamations such as ‘Down with cheating!’ (in fact early versions of the theory were very explicit in making such comparisons), but this makes it hard to understand how we can use such statements in the range of ways that we do, for example, describing them as true or false, embedding them within longer statements, and treating them as premises and conclusions in arguments. In Chapter Three, I show why expressivism can indeed permit such practices.

It may also be thought that there are more convincing theories of normative and moral language than expressivism, and in Chapter Five, I attempt to show that this is not in fact the case by examining alternative theories and their shortcomings.

**1.2 Mixed normative statements**

The category of normative statements includes *mixed* normative statements, which function partly by expressing attitudes, but also assert descriptive propositions, i.e., propositions that could equally well be asserted merely by making descriptive statements. One way that mixing can occur is through explicitly combining normative and descriptive statements, as in:

 LAWN

The lawn is in the back garden, which is the best arrangement.

LAWN is a mixed statement, as it contains a normative statement after the comma, but also asserts a descriptive proposition before it. As well as expressing an attitude, it also expresses the speaker’s belief that a certain lawn is in a certain back garden. (I will call normative statements that are not mixed *pure*.)

But a more interesting way to make a mixed statement is to use so-called ‘thick’ normative terms, as in:

 HITLER

Hitler was cruel

which asserts a descriptive proposition about Hitler (that he thought and behaved in certain ways), as well as expressing an attitude towards that fact, a highly negative one. We can define a thick normative term, such as ‘cruel’, as a term which is normative but also has a descriptive element to its meaning. But note that the attitudinal element in a statement employing thick normative terms (a thick normative statement) is usually *cancellable*. In this case, one could intelligibly, if repellently, say ‘Hitler was cruel, but I don’t mind at all about that – I think his cruelty was justified’. A thin normative term is a normative term which does not have any such descriptive element to its meaning (thin terms include ‘good’, ‘bad’, ‘better’, ‘worse’, ‘right’ and ‘wrong’). Admittedly, the idea that the meaning of a thick normative term can be separated into a descriptive and a normative component has been questioned by some philosophers.[[8]](#footnote-8) One possible misunderstanding can be cleared up by pointing out that those who, like me, think that there is a valid distinction between descriptive and normative content in thick normative meanings are not committed to the idea that there is a sharp boundary between cases in which a certain attitude is conveyed by the meaning of a statement and those in which it is inferred by listeners using extralinguistic knowledge. Consider this statement from a Wikipedia article about Russian settlement in the Philippines in 1949:

For months, the sleepy village [of Tubabao] became the sanctuary of people escaping the restrictions of communism, until they could be admitted to countries such as the United States, Australia, and France. (<https://en.wikipedia.org/wiki/Russian_settlement_in_the_Philippines>, retrieved 23.04.2022)

Wikipedia editors are required to write ‘from a neutral point of view’, which, on a very strict interpretation, could be taken to mean that all their statements should be purely descriptive, i.e., lacking all normative content. Now consider the characterisation of the immigrants as ‘escaping the restrictions of communism’. Does the use of this phrase carry any normative content? It is hard to give a definite answer to this.[[9]](#footnote-9) When someone uses the phrase ‘restrictions of communism’, there is quite a high chance that they are expressing a least a mildly negative attitude to the restrictions being referred to. But it is far from clear that the probability is so high as to justify the claim that this attitude is part of the *meaning* of this phrase. However, this is not a problem for our *general* ability to distinguish between the descriptive and the normative content of statements any more than the existence of sand patches that are barely moist is a problem for our ability to view the surface area of the Earth as being divided into land and sea.

As the above example shows, the term ‘normative content’ – like perhaps all terms outside mathematics, logic or theoretical science – is vague: it does not have sharp boundaries. Likewise for the thick terms themselves. ‘Cruel’, for example, lacks sharp boundaries for its extension and in this way resembles descriptive words that are vague, such as ‘red’. We could wonder, for example, whether a certain action was cruel or merely inconsiderate. This could affect the strength with which we criticise or condemn the action. In other cases we might disagree about whether the action was ‘cruel’ because we disagree about whether it should be criticised at all. (Perhaps, though it hurt someone, doing so was justifiable under the circumstances.) In such cases, our attitudes cause us to locate the boundary between the cruel and the non-cruel in different places.[[10]](#footnote-10)

Simon Blackburn has argued that, though thick normative terms do exist, the phenomenon is much rarer than the standard view (the one expressed in the preceding paragraphs) would suggest. He notes that if a house was described as having south-facing windows, this might cause a favourable attitude towards it, not because the phrase ‘has south-facing windows’ has some attitudinal component to its meaning, but only because of the contingent fact that such a feature tends to enhance the desirability of the house in the minds of those considering it. So it is, Blackburn suggests, with words like ‘courage’ and ‘temperance’: the positive emotional colouring possessed by these words is due to the fact that ‘exercises of courage and temperance are normally beneficial’ (Blackburn 2010, 131), and is therefore not a function of their meanings.

But this is simply wrong. It seems as if Blackburn has been misled by an instance of the very concept of thickness whose application he is trying to limit. It is easy to accept the claim that courage is normally beneficial, but if we replace the word ‘courage’ by some more neutral term such as ‘deliberately acting in disregard of danger’, then the resulting claim becomes much more questionable (it is questionable, in other words, that deliberately acting in disregard of danger is beneficial more often than not). The attitudinal aspect therefore has to be carried by the meaning of the word. And, indeed, if I say that Anita was courageous, and you register what I say without knowing anything else about Anita, or about the context of my remark, you are entitled to think that I approve of whatever it was that Anita did.

I have conceded that the attitudinal component of the meaning of a thick normative term can be cancelled by the speaker. Actually, the cancellability property applies, not only to thick terms, but to most thin terms as well.[[11]](#footnote-11) If I say ‘This is a good knife’, my meaning is given by the fact that, subject to a certain qualification, I am thereby expressing a positive attitude to the knife *qua* knife. The qualification is that if I follow this remark with something like ‘In saying this, I don’t mean to imply that I approve of it in any way – I just mean that it cuts effectively and has no properties which would normally be thought bad in a knife, but in fact I don’t particularly approve of knives with these attributes’, then what I have said may be eccentric, but it is not logically incoherent. It is as if I had used ‘good’ in inverted commas.

**1.3 The Redundancy Thesis**

I maintain that there is nothing more to normativity than the expression of attitudes. Since attitudes are by there very nature subjective, accepting this view makes me a kind of subjectivist, though, not of course, a speaker subjectivist. I call this position ‘subjectivism about norms’ or ‘normative subjectivism’.

My theory of norms is also *anti-realist*. A good way to explain its anti-realist character is through the following claim:

 **Redundancy Thesis (RT)**

 If we did not make any normative statements, but instead only described things non-normatively (with reference, where necessary, to our own subjective attitudes), the only losses, in general, would be, at most, stylistic variety, rhetorical force, elegance and/or succinctness.

As earlier emphasised, expressivism is not the same as speaker subjectivism: expressing an attitude is not the same as describing or asserting it. However, according to the RT, we could in principle dispense with the making of normative statements in favour of simply describing our attitudes, and the result would be that our discourse would be diminished only in respect of the ‘surface’ qualities mentioned in the RT.

To illustrate the application of the RT, here is part of an editorial from a magazine containing a number of normative statements, followed by a paraphrase in which the normative statements have been replaced by statements describing the attitudes of the author(s).

For the past 30 years, it has been possible to believe that the risk of world war had receded and that the perilous days of nuclear brinkmanship belonged to an earlier era. The Soviet Union had collapsed. China was opening up. Liberal democracy was in the ascendent and the days of brutal, reclusive dictators appeared to be numbered. But now we must confront the world as it really is, and comprehend the crisis with which we are faced. The West must redouble its support - both military and economic - for Ukraine, and as a signal to authoritarian regimes elsewhere. We must meet tyranny with resolve. (From *The New Statesman* editorial, 14-20 October 2022, p. 3.)

Here is my non-normative paraphrase:

For the past 30 years it has been possible to believe the probability of world war had diminished and that the days of nuclear brinkmanship, in which the probability of things that we greatly fear happening, belonged to an earlier era. The Soviet Union had collapsed. China was starting to interact more with the rest of the world. Liberal democracy was becoming more common and the days of reclusive dictators causing extreme harm for their citizens without what we would consider to be adequate justification appeared to be numbered. But now we categorically[[12]](#footnote-12) want people to confront the world as it really is, and comprehend the climax of our difficulties. We categorically want the West to redouble its support – both military and economic - for Ukraine, and as a signal to non-democratic systems elsewhere. We categorically want to respond to the existence of political systems in which the executive exerts control by causing fear in those who oppose it with a strong determination to oppose such regimes.

Some might question whether this paraphrase really abolishes all hint of normativity. In assessing this, it is important to realise that an absence of normativity does not necessarily deprive a piece of discourse of all emotive force. For example, the description of a political system as one in which ‘the executive exerts control by causing fear in those who oppose it’ is likely to strike people with good democratic attitudes as an unattractive characterisation, but it is, for all that, a pretty neutral description of what tyranny consists in.

It would also be true to say that a reader of the paraphrase would be as well informed about the author’s attitudes concerning the subject-matter as a reader of the original would be. Such is the emotive effect of the frequent repetition of the phrase ‘categorically want’. But this does not alter the fact that this expression, as well as all the others used, is descriptive, not normative.

The prose of the new version is of course extremely leaden, because it needs to use more words in order to avoid some of the normative expressions in the original. But that is not a problem for defenders of the RT. It is not being claimed that we would be better off if we always avoided normative expressions. Indeed, the principle itself lists the main things we would lose out on if we did this, and these include elegance and succinctness.

The RT is a strong claim which is not obviously true. Thus Michael Ridge maintains that the fact that there are reasonable translations of English words such as ‘good’, ‘ought’, ‘reason’ and ‘must’ into all other natural languages ‘reflects something philosophically deep about our nature as reflective thinkers and deliberators’ (Ridge 2015, 9). This is probably a common view, but I suggest that it is false. Our fundamental nature as reflective thinkers and deliberators would be untouched in the (admittedly improbable) event that we were to stop using normative statements. And Ridge’s claim is an interesting falsehood because it shows that although *speaker* subjectivism is not a true account of normative statements, it is not so very far from the truth as is often supposed. Bear in mind that the RT does *not* say that normative statements can be ‘reduced’ to non-normative statements which are analytically equivalent in meaning. The requirement is only that the basic intent of the original should be preserved. For this reason, cases in which non-assertoric normative statements are embedded within longer statements, which have been so vexatious for advocates of non-cognitivist theorists of normativity (and which will occupy our attention in Chapter Three) need not trouble the defender of the RT. Consider, for example:

If nuclear war is wrong, then we should protest against the installation of the missiles here.

Although, strictly speaking, it does not mean the same thing, the *intent* behind this remark could often be expressed by:

Because we think nuclear war is wrong, we are in favour of protesting against the installation of the missiles here.

Those who are not willing to accept the RT face a challenge: Present a piece of discourse in which the normative words cannot be replaced by purely descriptive ones without the loss of something more weighty than stylistic variety, rhetorical force, elegance and/or succinctness. I believe this is a challenge which cannot be met.

**CHAPTER TWO**

**Expressivism: The Basics**

**2.1 Expressing attitudes**

In this chapter, I will explain expressivism in more detail, though the full elaboration of the theory will be presented in Chapter Three. As explained in the last chapter, according to expressivism, simple, stand-alone normative statements do not assert that the speaker has certain attitudes; rather they express those attitudes. But what does that actually mean?

It is generally required or demanded that a speaker does not state something unless they believe it to be true. That is to say, they are generally open to criticism if they do not have the relevant belief. Mark Schroeder calls this an ‘assertability condition’ for a sentence and understands it as ‘the condition under which it is semantically permissible for a speaker to assert it’ (Schroeder 2008, 31). There are apparent exceptions to it such as joking or play-acting, but perhaps in such cases we should not say that a genuine *assertion* is being made. Understood in this way, the requirement seems to be almost tautological – it is part of the meaning of the word ‘assertion’ that in making an assertion of a descriptive statement, you open yourself to criticism if you do not believe what you have asserted. If a non-tautological formulation is preferred, we might say that if *under certain circumstances* (such that you are not signalling, and the context does not indicate, that you are joking or play-acting etc.), you utter a sentence of a certain grammatical form, there is a certain belief (which one obviously depending on the words in the sentence, and often the context as well) such that you are open to criticism if you do not hold it.[[13]](#footnote-13) The act of assertion, as well as the statement (token) asserted, may be said to express the corresponding belief.

Expressivists explain the sense in which a stand-alone normative statement expresses an attitude in a similar way. For example, ‘Honesty is good’ is assertable only if the speaker has a positive attitude to honesty. If they do not have this attitude, they are generally open to criticism for having made the statement. In general, any normative statement is assertable only if the speaker has a certain attitude.[[14]](#footnote-14) The statement (token) asserted may be said to express that attitude.

An expressivist may want to treat this analogy as involving a *contrast* between descriptive and normative statements. That is to say, they may want to claim that normative statements express attitudes *instead of* beliefs. This is the traditional, ‘radical’ expressivist view, of which an example is A.J. Ayer’s ‘emotive theory’ (Ayer 1946, 102-114). Alternatively, they may prefer to say that *all* statements express beliefs, so that normative statements express both beliefs and attitudes. This might be considered the ‘common sense’ view, as we do habitually speak of such things as the ‘belief that euthanasia is wrong’. It is associated with more recent forms of expressivism such as Simon Blackburn’s ‘quasi-realism’ (Blackburn 1984, 181-223; 1998). This point will be further discussed in Section 3.3.

What exactly justifies the analogy between the attitude requirement for normative statements and the belief requirement for descriptive statements (or for all statements)? This can be located in the idea of *sincerity*. At first, it might be thought that the link is simply that if the speaker does not have the relevant belief or attitude, then they are being insincere. However, Michael Ridge has pointed out that the situation is more subtle than this. Ridge provides the following imaginary example:

Consider first a Freudian case in which Bob believes that he believes his mother loves him but actually does not believe that she loves him. In fact, Bob believes his mother hates him. There are familiar ways in which Bob’s behaviour might show that he has deluded himself about his own beliefs. For example, he might predict that his mother will do things that would make sense only if he believed she hated him. Perhaps Bob also feels a sort of anxiety people associate with being hated by a close family member when the subject of his mother arises, goes out of his way to try to please her and is highly deferential to her, etc. Nonetheless, because he cannot cope with the idea that his mother hates him he has somehow convinced himself that he believes that his mother loves him. (Ridge 2006, 488-9.)

Another type of case mentioned by Ridge is taken from real life. Drawing on the work of Gregory Currie, he cites the case of certain schizophrenics who can be interpreted as being unable to distinguish between their beliefs and their acts of imagination (Ridge 2006, 489-490). For example, they may claim to have paranoid beliefs, but fail to act on them, thus supporting the contention that they do not really believe what they say they believe. It is therefore perfectly possible for a speaker to make a descriptive statement sincerely and yet not have the corresponding belief. Ridge concludes on the strength of these cases that the condition for a speaker to be sincere is not that they have the relevant belief but that they *believe* they have it. The same thing, he maintains, seems to apply to normative statements and attitudes. There could be cases analogous to those just described in which the person concerned is mistaken about the attitude they have, but that does not in itself make them insincere. They are sincere provided they *believe* they have the relevant attitude. As we noted above, if people do not have the attitudes which their statements express, they are generally open to criticism, though we can see now that this would not necessarily be for insincerity. What we might say in such cases is merely that they ‘misspoke’.[[15]](#footnote-15) (Of course, in practice, since cases in which insincerity and ‘misspeaking’ come apart are quite rare in this context, it is likely that they *are* being insincere.)

Another reason for thinking that attitudes stand to normative statements as beliefs stand to descriptive statements can be gleaned from consideration of *inferences*. Suppose I make the inference ‘Deception is wrong. Simon deceived me. Therefore Simon did something wrong’. Then I acquire a new attitude, one concerning Simon’s action. This seems to parallel the fact that if I make the inference ‘Any solution which turns litmus paper red is acidic. This solution turns litmus paper red. Therefore this solution is acidic’, I acquire a new belief concerning the solution referred to in the inference.

**2.2 Geach on ‘good’**

Two of the most common normative words are ‘good’, and its negative counterpart ‘bad’. According to P.T. Geach, ‘good’ is an *attributive*, as opposed to a *predicative* adjective (Geach 1967, 64). To illustrate the concept of an attributive adjective using one of Geach’s own examples, the adjective ‘big’ in ‘big flea’ is said to be attributive, meaning that the adjective and noun cannot be semantically separated: a big flea is not necessarily a thing which is both big and a flea (for otherwise, we could easily argue that a big flea must be a big animal); whereas, in the phrase ‘red car’, ‘red’ is said to be predicative, meaning that such separation is possible: a red car is necessarily something which is both red and a car.[[16]](#footnote-16) In general, for an adjective ‘*a*’ to be attributive in Geach’s sense, for any noun ‘*n*’, being an *a n* must not entail being both *a* and an *n*. The adjective ‘good’ is attributive in this sense, because, for any noun ‘*n*’, being a good *n* does not entail being both good and an *n*.

Geach says that ‘it belongs to the *ratio* of ‘want’, ‘choose’, ‘good’, and ‘bad’, that, normally, and other things being equal, a man who wants an *A* … will choose an *A* that he thinks good and will not choose an *A* that he thinks bad’ (ibid., 69). This suggests it is impossible to know whether anything is good unless one knows in relation to which general set of things its goodness is to be assessed. Geach maintains that the meaning of ‘good’ is actually descriptive, even though the particular descriptive criteria to be applied vary from case to case, depending on what specific noun is involved. And while he accepts that it is possible to say ‘*x* is good’ without using any noun, he thinks that some noun or other must be implied by the context for us to be able to make any sense of the statement (ibid., 65-6). Not that any noun will do, on his view. For example, the phrase ‘good event’ is senseless, as ‘event’ is ‘too empty a word to convey … a standard of goodness’ (ibid., 71-2).

In Chapter One, it was claimed that If I say ‘This is a good knife’, my meaning is given by the fact that I am expressing a positive attitude to the knife *qua* knife (subject to the possibility of cancelling this expression of attitude). The inclusion of the phrase ‘*qua* knife’ is a clear point of contact between this analysis and Geach’s - it tells us in relation to which general set of things the goodness of the knife is to be assessed. According to my use of the word ‘descriptive’, which would rule out the statement’s expressing any attitude, this is no reason to call such statements descriptive, though that might just be considered a minor terminological difference. (According to my terminology, ‘This is a good knife’ is a mixed normative statement, asserting the descriptive fact that the object is a knife, as well as expressing a positive attitude towards it *qua* knife).

What should be said about Geach’s claim that there is a conceptual connection between ‘good’ and acts of choosing to the effect that a person who wants an *x* will, all other things being equal, choose a good *x* as opposed to one which is not good? (This is how I interpret his claim that this principle is part of the ‘*ratio*’ of the words involved.) It does seem to work well for knives and other objects which serve some specific purpose. But for people, ‘good’ and ‘bad’ usually work differently. More often than not, ‘She is a good person’ means that the person concerned is *morally* good. It might be said that you would choose such a person to rely on to perform morally good actions and avoid morally bad ones, but this seems very contrived. It seems better to omit the reference to choice and say that a morally good person just *is* someone who generally performs morally good actions and avoids morally bad ones (the force of the word ‘morally’ will be discussed in Chapter Four).

Geach seems to be right to say that ‘good’ is always an attributive adjective. But then how are we to make sense of the simple statement ‘John is good’? The answer, as so often in matters relating to language, is provided by reference to the context of utterance. John could be many things: for example, a cricketer, an inhabitant of the UK or just a person (in which case the moral criterion normally becomes applicable). He could be evaluated *qua* any of these, and context will determine which is intended.

But what exactly is it to evaluate a thing *qua* something or other? Imagine an eccentric person who collects toasters and values them only for their aesthetic properties. For him, approving of a toaster *qua* toaster is simply a matter of finding it aesthetically appealing. If he asserts ‘This is a good toaster’ purely on the basis of its aesthetic qualities and not for more conventional reasons (such as its ability to toast bread safely and effectively), has he misused the word ‘good’? It seems to me that he hasn’t, and so I would suggest that in making the statement ‘x is a good *n*’, a speaker expresses a positive attitude to *x* using whatever criteria he prefers for evaluating *n*s. Of course, it is worth pointing out that where the criteria used are very different from those which most people would use, a cooperative language-user would draw attention to this difference.

Finally, what about Geach’s claim that the variable in ‘good *n*’ could not be taken by any noun whatsoever? If I am right that individual criteria of evaluation, as opposed to socially agreed ones, can be acceptable, we should probably reject this. To take Geach’s actual example, even if there are no socially agreed criteria for evaluating events, the speaker could appeal to their own. If I say ‘This year’s Notting Hill Carnival was a good event’, I could just have some particular criteria of my own in mind, which I could clarify if requested. Incidentally, there is an interesting puzzle about ‘event’ in this regard. A pretty good definition (that is, good *qua* definition!) of ‘event’ would be ‘something that happens’, and yet ‘This year’s Notting Hill Carnival was a good thing to happen’ is much clearer than ‘This year’s Notting Hill Carnival was a good event’ in terms of how the carnival is meant to be evaluated. The former seems to focus attention on consequences, whereas the latter lacks this focus, and indeed the basis of the evaluation might be entirely unclear absent any explanation from the speaker. Why would this be the case if an event just *is* a thing that happens? I think the answer is that it is not just meanings that matter here but also mental associations created by known patterns of usage. ‘Good thing to happen’ is a phrase which is often used when the consequences of an event are being considered and this creates an expectation that it is consequences which are being appealed to when this phrase is applied to Notting Hill Carnival, an expectation which is absent when it is merely described as a good *event*.

**2.3 Deontic words**

Having considered the evaluative word ‘good’, it is now time to investigate the set of normative words commonly referred to as ‘deontic’, namely, ‘right’, ‘wrong’, ‘obligatory’, ‘permissible’, ‘ought’, and so on. What all such words have in common is that they apply to actions. Focusing initially on ‘ought’, to say of an action that it ought to be done is to express a positive action towards the doing of that action.[[17]](#footnote-17) However, a more precise account requires distinguishing between types of action and individual actions. If an individual *i* says at *t*, for example, that one ought to help injured people, then part of what is involved in the positive attitude which *i* thereby expresses at *t* is a tendency for *i* to help injured people, all other things being equal. This has been referred to as the *motivational load* of deontic statements (Sayre-McCord 2021, #3).[[18]](#footnote-18) But bear in mind that *i* is saying, not that they specifically ought to help injured people but that *one* ought to help them. To reflect this fact, we should also add that in such cases there will be a tendency not only for the speaker to carry out such actions, but also to advise others to carry them out. (This in effect makes the present theory of deontic statements a form of *prescriptivism*). In general, when *i* says at *t* that one ought to *x* (for some action-type *x*), they have mis-spoken if they do not have a tendency to *x* and to advise others to *x* in cases where they believe these others would follow their advice.

‘Ought’ statements can be merely prima facie or ‘decisive’. The difference lies in what may defeat the tendencies identified in the previous paragraph. When the ‘ought’ statement is prima facie, it is sometimes defeated when the speaker would accept a statement of the form ‘One ought to *y*’, and doing *y* would be incompatible with doing *x* under the circumstances. When, on the other hand, it is decisive, such defeat is not possible by definition. This is not to say that defeat is never possible in the case of decisive ‘ought’ statements. For even if *i* thinks that one ought, decisively, to do *x*, they might fail to do *x* merely through weakness of will. Note that a negative attitude towards an action-type can itself be described as either prima facie or decisive, depending on the above defeasibility consideration. Indeed, whether a belief in the wrongness of an action-type is prima facie or decisive depends on whether the corresponding attitude is prima facie or decisive.

When an ‘ought’ statement applies to an individual action, as opposed to a type of action, the connection with the speaker’s behaviour is more indirect. If I say ‘Simon ought not to have done that’ without any indication of why I disapprove of Simon’s action, then I do not thereby commit myself to behaving in any particular way or giving any particular advice to others. However, people approve or disapprove of individual actions for reasons. The reason always lies in the fact that the action belongs to a general type which the speaker would approve or disapprove of (where a full description of the type might have to include the circumstances in which it is or would be done and/or its consequences).[[19]](#footnote-19) It is this more general attitude which has the behavioural implications we have noted. (It might be thought that by saying Simon’s action was wrong, one would at least by implying that one would avoid, and advise against, *similar* actions. But similar in what respect? The relevant action-type provides the answer to that question.)

We can build on what we have said about ‘ought’ to explain the other deontic words mentioned above. Roughly speaking, an action[[20]](#footnote-20) is said to be wrong when the agent ought not to do it (or have done it). (Here the ‘ought’ may be understood in a prima facieor decisive sense, giving correspondingly different senses for the word ‘wrong’.) An action is said to be the right action when it is the only action possible under the circumstances that would not be wrong. To say that an action is obligatory is to say with particular force that it *alone* is the right action. To say that an action is permissible is to say that it is not the case that it is a wrong action.

This theory lays great emphasis on behaviour. But there is of course another aspect to having an attitude, which is a tendency to be subject to various emotions (the phenomenological aspect). For example, thinking or saying ‘We had a good time that day’ is likely to be accompanied by pleasant feelings. On the other hand, this phenomenological aspect of attitudes is very variable. Consider the following moral example. When uttering ‘Deceiving people is wrong’, one person might feel a significant level of negative emotion, while another may not. But if the latter person would, at least when all other things are equal, avoid deceiving people, and, when the opportunity arises, advise others against deception, then they can, I think, count as adhering to the assertability condition for this statement, according to which the speaker must have a negative attitude to deception. This is why the treatment of attitudes in my theory places such heavy emphasis on actions, including the action of giving advice. But it does not deny the phenomenological aspect of attitudes.[[21]](#footnote-21)

Deontic words and evaluative words, though they form distinct categories of normative words, are of course linked in meaning. The link is provided indirectly by Geach’s observation, mentioned earlier, that a person who wants an *x* will, all other things being equal, choose a good *x* as opposed to one which is not good. This observation indirectly indicates a link between the evaluative word ‘good’ and the deontic word ‘ought’, because any speaker who was disposed to make such a choice would accept that, prima facie, he *ought* so to choose.

**2.4 Speech and thought**

Up to now, I have talked almost exclusively of normative *language*. But there is clearly also such a thing as normative *thinking*, as when I think privately to myself that bribery is wrong or that I gave the wrong answer to an arithmetic question. Should expressivists claim that their theory applies to normative thought as well as normative language?

I believe the answer is yes. Recall that the distinctive claim of expressivism is that each normative statement, when used stand-alone, expresses an attitude, which is another way of saying that its utterance is subject to an assertability condition, according to which it should not be uttered unless the speaker has the relevant attitude. In the case where someone is not uttering the statement, but merely thinking it, it seems reasonable to suppose that there is a counterpart to this assertability condition applicable to thoughts, according to which the person will not think it correct to have the relevant normative thought unless they have the relevant attitude.

In adopting this theory, we are treating thought as a kind of inner monologue subject to similar constraints as ordinary speech. There are two objections that might be made to this. One is that phenomenologically speaking, not every thought that we have takes the form of inner speech: sometimes a thought comes to us in a non-linguistic form, often because we cannot at that moment think of the appropriate words. Another, related, objection is based on the fact that some non-human creatures may have normative thoughts (such as that it would be better to go and hunt now than to sleep). Given that these creatures do not have language in the ordinary sense, they surely do not have any form of inner language serving as the vehicle for their thoughts.[[22]](#footnote-22) Against this, one might doubt it is possible to draw the distinction between a creature’s thinking that it would be better to go and hunt and thinking merely that they want to go and hunt outside of the conventions of something pretty close to a human language. In any case, I think the best response to this objection, and also to the previous one, is to retreat to a very modest position, which is simply that even if not every normative thought (of a person or non-human animal) is expressed in language, there is bound to exist a possible linguistic expression of it, and for this reason, it is indeed possible to regard the present theory of normative speech as applying to normative thought. However, I would also claim that normative thought, like its linguistic counterpart, is redundant in the sense that it does not add anything of substance to a person’s awareness that they have certain attitudes and desires, for it would be highly implausible to accept redundancy for normative speech but deny it for normative thought.

**CHAPTER THREE**

**Expressivism: Truth and the Embedding Problem**

**3.1 Geach’s challenge**

Expressivism has long been subject to a challenge from the philosophy of language which has come to be known as the ‘Frege/Geach point’. Geach explains it using the following example:

TORMENTING

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. (Geach 1965, 463)

His comment about this argument is as follows:

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, 463-4.)

It may seem as if there is an obvious way in which expressivists can avoid the force of Geach’s argument. It is true that in the first premise, ‘bad’ is not being used in the way specified by the expressivist--it is not being used to condemn anything (either in the antecedent or the consequent of that premise). But the expressivist’s claim that normative statements express attitudes should be limited to stand-alone statements like the second premise and the conclusion, i.e., cases in which the normative statement is not embedded within a longer statement.[[23]](#footnote-23) Furthermore, while it is true that there is a shift in the way ‘bad’ is used in the course of the argument, this fact alone cannot be used to show that it is invalid by the expressivist’s lights. For if it were, we would have not merely a point against expressivism, but a paradox to be addressed by anyone who (rightly) thinks the argument valid.

To see this, consider the following purely descriptive analogue of Geach’s argument:

 JUGGLING

If doing a thing is hard, doing it with your eyes shut is hard.

Juggling is hard

 Ergo, juggling with your eyes shut is hard.

In the first premise, the descriptive word ‘hard’ is not used to express the belief that anything is hard. But it is so used in the second premise and the conclusion. Yet we would not say that this shift in meaning (if one wants to call it that) renders the argument invalid. As we have seen, the expression of beliefs by descriptive statements is analogous to the expression of attitudes by normative statements. The two arguments show that we need to confine this conception of what each type of statement expresses to stand-alone cases. Once this point is understood, there are no grounds for supposing that the expressivist would be forced to regard the first argument as involving the fallacy of equivocation.

But there is an important question left unanswered by these remarks, which is: What exactly is the word ‘bad’ doing in its two occurrences in the first premise of Geach’s argument if it is not expressing an attitude? The use of ‘hard’ in the corresponding positions of JUGGLING may not seem so problematic, as we have logical theory to turn to in understanding the enveloping statement. The latter seems roughly equivalent to:

HARDNESS

 (ꓯ*x*)(*x* is hard → doing *x* with your eyes shut is hard),

which says that there are no cases of a thing’s being hard such that doing it with your eyes shut is not also hard. Here it is not a problem that neither component is believed. ‘Hard’ still seems to have its normal descriptive use.

We say that an argument is deductively valid when its conclusion is logically guaranteed to be true if the premises are true.[[24]](#footnote-24) This suggests that to talk about the validity of Geach’s argument, we need to have some sort of conception of truth for normative statements. Their grammatical form as *statements* would also seem to require this. But how does this square with the expressivist’s view of them as expressing attitudes in their stand-alone uses? Allied to the question of truth, there is also the issue of falsity and the use of the word ‘not’ in normative statements. We might suggest that saying something is not wrong expresses the absence of a negative attitude towards it. But this can’t be the whole story about negated normative claims, because the reason for the absence of disapproval might be merely that the speaker has never considered the issue, in which case they would not necessarily want to assert that the thing was not wrong. Thinking that something is not wrong is not simply the same as not thinking it wrong.

**3.2 Inferentialism**

But perhaps the key to understanding all this, is not, after all, a matter of obtaining a theory of the truth and falsity of normative statements. There is an attractively simpler alternative. While we can agree that deductively valid arguments are necessarily truth-preserving, we may think that there is a more insightful way of explaining why some arguments are deductively valid while others are not. This depends on the properties of logical operators such as ‘if … then’, more specifically on the ways in which their use licenses inferences from certain stand-alone statements to others (e.g., ‘If *p* then *q*’ licenses an inference from *p* to *q[[25]](#footnote-25)*). When such statements happen to be normative, expressivism tells us how to understand them: as expressions of attitude. Contrary to the suggestion of the previous section, we do not need, in addition to this, an account of what normative words mean in non-assertoric embedded contexts. All the work is done by our account of the inferential properties of the operators, which is the same irrespective of whether the embedded statements are normative or descriptive.[[26]](#footnote-26) The central idea, known as *inferentialism*, is that an understanding of the meaning of words is to be derived from an understanding of their inferential role in the language.[[27]](#footnote-27) Closely associated with inferentialism is the *minimalist theory of truth*, according to which (on one formulation, at any rate) to think that *p* is true is merely equivalent to accepting that *p*.[[28]](#footnote-28) On this view, truth-preservation in valid arguments is nothing more than the guarantee that if we can accept the premises, then we can also accept the conclusion. As well as adopting minimalism about truth, we can also affirm minimalism about truth-*aptness*, according to which it is not necessary to place any restrictions on the kinds of sentences which can be true or false other than the requirement that they should have a certain grammatical structure. As far as English is concerned, this amounts to their ability to be embedded in ‘that’-clauses.[[29]](#footnote-29) Notice how this form of expressivism diverges from the earlier ‘emotive theory of ethics’ (as, for example, in Ayer 1946, 102-114) in allowing normative statements to count as true or false. As Paul Horwich puts it, albeit using slightly different terminology from that adopted here, ‘expressivism may and, I think, should be located, not in the thesis that evaluations don't express beliefs and can't be true, but rather in the thesis that evaluative utterances express desires’ (Horwich 1994, 20).

There is one respect in which this account does not tell the whole story. It concerns negation. In fact, there are two aspects to this: negation in general, and the negation specifically of a normative statement. As far as the former is concerned, it has been argued that any convincing account of negation must include its connection, not simply with the acceptance of statements or propositions, but with their rejection. For example, if we accept ‘It is not the case that there is life on Mars’, then we should reject ‘There is life on Mars’. But in order to accommodate this, in addition to rules of inference concerning what statements we ought to accept, we should also have rules of inference concerning what statements we ought to reject (for example, ‘If *p* then *q*’ licenses an inference from the rejection of *q* to the rejection of *p*). So-called bivalent logics have been developed which place acceptance and rejection on an equal footing (Rumfitt 2000). As for the second aspect of the problem, that of understanding how *normative* statements can be negated, that will become apparent a little later.

That aside, is the inferentialist view generally viable? One sceptic is James Dreier (Dreier 1996). He imagines an interjection ‘hiyo’ which can be used to accost someone, e.g., ‘Hiyo, Bob!’. He then imagines using the word ‘hiyo’ in an indicative[[30]](#footnote-30) sentence, e.g., ‘Bob is hiyo’, in such a way that ‘Bob is hiyo’ and ‘Hiyo, Bob!’ are interchangeable, but points out that the idea of minimalist truth-aptness alone would not enable us to understand what this sentence meant in non-assertoric embedded contexts such as ‘If Bob is hiyo, then Jim is hiyo’. It is the same, he suggests, with a normative predicate like ‘wrong’: If ‘Capital punishment is wrong’ is understood in expressivist terms, then the idea of minimalist truth-aptness alone does not enable us to understand what this sentence means in non-assertoric embedded contexts.

But consider the following:

 If Bob is hiyo, then Jim is hiyo.

 Bob is hiyo.

 Therefore, Jim is hiyo.

Suppose Stella sees Bob and Jim coming up the street together. She decides to accost Bob in the Dreier style by saying ‘Bob is hiyo’. But suppose she happens to accept the conditional statement ‘If Bob is hiyo, then Jim is hiyo’. She remembers this and uses the above argument to draw the conclusion ‘Jim is hiyo’, which she then uses to accost Jim. In effect, the conditional licenses one speech-act on the basis of another.

Now of course in the English language as we actually have it, the above argument is gibberish. But what is notable about the story I have just told is that it does involve the first premise being used recognizably like a normal conditional (this of course entirely accords with the view of ‘if … then’ explained above as an operator whose role it is to mediate what can be inferred from what). And if such a story can be told for ‘hiyo’, why not also for normative words such as ‘wrong’?[[31]](#footnote-31)

But how, it might be asked, do conditionals like the first premise of TORMENTING get to be accepted in the first place? The key to this lies, I think, in the presence of causal connections in the acquisition of attitudes. Disapproving of something tends to bring with it disapproval of getting others to do it. Becoming conscious of this, you may then encode it, so to speak, in the form of a generalisation: ‘Whenever something is wrong, it is wrong to get someone to do it’. Even if you do not do this explicitly, you may be said to accept it implicitly if you would tend to accept the first premise of TORMENTING and other relevantly similar conditionals.[[32]](#footnote-32)

Having a certain attitude, as well as causing a person to have an attitude, can also cause or be caused by the holding of a descriptive belief. Thus learning that someone was a slave trader may tend to cause Amy to want his statue removed and she can, if she likes, codify this in the form of the generalised conditional ‘If someone is a slave trader, then there should not be a statue of them’.[[33]](#footnote-33)

Though it is less common, the causal connection can go the other way too. This arises from the fact that there are only a finite range of reasons why a person might think a certain normative attribute such as wrongness would be applicable to something. Furthermore, in a given context, the range may be very small indeed, simply because that context rules out many possibilities. In such a case, the individual may be able to accept:

WRONGNESS

For any *x*, if *x* is wrong, then *x* is *f*

where ‘*f*’ is descriptive. This could mediate an inference from ‘*a* is wrong’ (for some *a*) to ‘*a* is *f*’, so that my believing that *a* is wrong causes me to believe that *a* has a certain descriptive property. (This might happen because I have been told by someone with relevantly similar attitudes to me that *a* is wrong.)

We can now see, I think, how an inferentialist might try to handle the negation of a normative statement. Suppose I believe the following:

SUFFERING

If action *a* is wrong, then *a* causes suffering.

Now suppose I also come to believe that *a* does not cause suffering. Then of course, by the modus tollens inference rule, I can infer that *a* is not wrong. We can have negated normative statements partly because we can have conditionals like SUFFERING.

But there are also other kinds of conditionals that help to make sense of negated normative statements. Thus suppose that if I believe someone is a racist, then I have a negative attitude towards them, and consequently I accept:

RACIST1

For any *x*, if *x* is a racist, then *x* is bad.

But regardless of any specific connections between goodness and badness and descriptive facts, ‘good’ and ‘bad’ are incompatible predicates, since I cannot have both a positive attitude and a negative attitude to the same thing (assuming that we are talking about overall attitudes and not prima facie ones). Hence, if I accept RACIST1, I should also accept

RACIST2

For any *x*, if *x* is a racist, then *x* is not good

which could mediate an inference from a descriptive statement (one which attributes to some person the descriptive attributes of a racist) to a negated normative statement of the form ‘*x* is not good’.

The device of negation also brings with it the notion of falsity, since it is false that *p* if and only if it is not the case that *p* and the statement ‘*p*’ is false if and only if it is not the case that *p*. Hence we have a notion of falsity applicable to normative statements.

The inferentialist approach seems to take care of a wide range of cases in which normative statements are embedded within longer statements. In addition to ‘if … then’ and ‘not’, which we have already considered, we might cite the following:

1. Either he acts rightly or he will face the consequences.
2. Possibly, he has acted rightly.
3. Their action will make the outcome bad.
4. All lies are wrong.
5. Most lies are wrong.

Of these, (1) and (4) are the most straightforward. The logical behaviour of the words ‘or’ and ‘all’ are well understood, and there is no reason to think that the fact that one of the disjuncts of (1) is normative affects the correct understanding of (1), or the fact that ‘All lies are wrong’ is normative affects the correct understanding of (4).

In understanding (2), recall that whether we have a certain attitude to something can be affected by other beliefs we have, including descriptive beliefs. Such links can vary in strength. Sometimes they compel belief, while on other occasions the link is less strong. (2) can be analysed as: ‘There is no true proposition which compels the belief that he has not acted rightly’. This condition is understandable in expressivist terms, because the belief that an action was not right is understandable in expressivist terms. For we saw above how, on the expressivist view, such negative normative beliefs can arise.

It might be objected to the above analysis that people who are capable of understanding and believing (2) do not necessarily understand the phrase ‘there is no true proposition which compels the belief that he has not acted rightly’ and hence that a belief in (2) cannot be analysed as being equivalent to the belief conveyed by this phrase. However, it is not uncommon for a philosophical analysis of an expression to contain concepts or ideas which the users of the expression would not necessarily understand. It is legitimate provided we understand that when it is claimed that such users have the relevant belief, the word ‘belief’ is understood in a special sense. It does not imply an explicit understanding of what is believed. The understanding required consists merely in the ability to *recognise* the condition as obtaining. The mental state involved is the same as that characterising someone who recognises that *i* and *j* are *first cousins* by looking at a family tree, even though they would not be able to state anything resembling the formal definition, ‘*x* and *y* are first cousins just in case their closest common ancestors are their grandparents’. It is useful to have a technical term for this mental state. I shall call it *effective* belief. To say that (2) is to be analysed as ‘There is no true proposition which compels the belief that he has not acted rightly’ is equivalent to saying that the assertability condition for (2) for *i* at *t* is that *i* effectively believes at *t* that there is no true proposition which compels the belief that the person referred to by ‘he’ in (2) has not acted rightly.

Now for statement (3). This makes sense because normative propositions can be caused to be true. This is the case because our attitudes to a thing depend on the descriptive properties we believe it to have, and a thing can be caused to have certain descriptive properties. Thus (3) might be true because the action of the people referred to will cause someone to have the descriptive property of undergoing suffering.

(5) means that in a sufficiently large and varied sample of lies, more than half of them will be wrong. This is understandable in expressivist terms, since the expressivist can explain what it is to believe an individual action to be wrong: it is to have a negative attitude towards it. This is *not* of course to imply that someone who says that most lies are wrong is simply asserting that they would disapprove of more than half of all lies in a sufficiently large and varied sample of lies. This would be like saying that if someone thinks that most swans are white, then it is true of more than half of all swans that they would believe that they were white. To say that more than half of some set of things have a certain property is not the same as saying that someone or other would believe them to have that property. Thus to say that more than half of all lies in a large and varied sample of lies would be wrong is not the same as saying that someone would believe them to be wrong, where believing something to be wrong is having a negative attitude towards it.

It might be argued that whereas we can readily distinguish between a swan’s being white and our believing that a swan is white, we cannot so readily distinguish between an action’s being wrong and our believing that it is wrong, if believing an action to be wrong is merely a matter of having a negative attitude towards it. But this is not so. In the former case, the distinction is possible because we may judge some of our beliefs to be false when assessed in relation to our current beliefs. Similarly, in the latter case, we may judge some of our attitudes to be mistaken when assessed in relation to our current attitudes. These attitudes which we judge to be mistaken can be equated to false normative beliefs on the expressivist view. This is all we need to make the distinction between an action’s being wrong and our believing it to be wrong.

So much for (1) to (5). There is, however, a class of expressions creating embeddings of normative statements which requires more careful discussion. These are verbs of propositional attitude, such as ‘believes’, ‘knows’, ‘wants’, ‘imagines’, ‘hopes’ and so on. They form the subject of the next section.

**3.3 Propositional attitudes**

The most important propositional attitude which can (it appears) be directed to a normative proposition is belief. In fact we have, of course, been assuming that normative beliefs exist, and that to believe a normative proposition is simply to have the corresponding attitude.[[34]](#footnote-34) For example, to believe that a certain action is wrong is to have a negative attitude to that action. This seems intuitively right, but here is an independent argument for it. The assertability condition for a given descriptive statement ‘*p*’ is that the speaker should believe that *p*. If this generalises to normative statements as well, then the assertability condition of, for example, ‘It is wrong to drink alcohol without restraint’ is that the speaker should believe that it is wrong to drink alcohol without restraint. But according to the expressivist, the assertability condition for ‘It is wrong to drink alcohol without restraint’ is that the speaker should have a negative attitude to drinking alcohol without restraint. This entails that the speaker’s believing the action to be wrong and having a negative attitude to the action are one and the same thing.

But *is* it the case that the principle that the assertability condition of a statement ‘*p*’ is that the speaker should believe that *p* applies to all statements including normative ones? It has been argued that whereas desires or attitudes have ‘world-to-mind fit’, beliefs have ‘mind-to-world fit’ (Ridge 2015, 6). In other words, we want the world to fit our attitudes, but we want our beliefs to fit the world. The idea of ‘fitting the world’ is philosophically vague, but I hope it is clear enough not to require any detailed elucidation here, which would take us too far from our central task. Notice that the statements we use to express our beliefs or attitudes could also be viewed in terms of the two types of fit. Since expressivists will say that normative statements express attitudes, they will say such statements have world-to-mind fit. But if an expressivist is also going to say that such statements express beliefs (as well as attitudes), it seems they will need to say that such statements have mind-to-world fit *as well*. But is this true?

Consider again the statement ‘It is wrong to drink alcohol without restraint’. Suppose we assent to this statement because we believe that drinking alcohol without restraint leads to suffering in the long run and we have a negative attitude to suffering. If drinking alcohol without restraint did not lead to long-term suffering, then (we can suppose) we would not have this negative attitude. There is here a genuine external constraint on our assenting to the statement, a mind-to-world fit. Thus to talk about our *believing* that it is wrong to drink alcohol without restraint seems entirely legitimate.

I will refer to the negative attitude to drinking alcohol without restraint, and the normative statement expressing it, as ‘non-fundamental’, meaning that the attitude is held because the person in question has some other attitude plus one or more factual beliefs[[35]](#footnote-35). In contrast with this, the negative attitude to suffering might be fundamental, as it might have no such basis. To put it another way, suffering may be considered intrinsically bad. In this case, there is no external constraint, in the relevant sense, on the person’s attitude or their making or assenting to the statement. The attitude comes, so to speak, ‘from the inside’ – it is not answerable to anything else. This is therefore not a case of mind-to-world fit.

Another example may help to convey the contrast. Suppose the statement is ‘Brutus’ action was wrong’, conveying a negative attitude to some action by Brutus. And suppose I assert this because I think that what Brutus did was to kill someone and I have a negative attitude to killing people. Then my attitude, and the statement expressing it, is non-fundamental. Mind-to-world fit seems to be involved here, as (according to my standards) my attitude fits what Brutus did. But my negative attitude to *killing people* may well be (for me) fundamental because it may well not be the result of my having some other attitude plus one or more factual beliefs. For me, in other words, killing may be intrinsically wrong, in which case my negative attitude to it does not need to fit some aspect of the way the world is.

Should we say then that fundamental normative statements do not express beliefs? I do not think we need to be as absolute as this. Instead, we can distinguish between a strict and a loose sense of ‘belief’. In the strict sense, requiring mind-to-world fit, such statements do not express beliefs (although non-fundamental normative statements do). But in a looser sense, every (stand-alone) utterance of a statement ‘*p*’, whether descriptive or normative, expresses a belief, and if ‘*p*’ is a pure normative statement, this belief is an attitude. (If it is a mixed normative statement, there is more than one belief expressed, at least one of which is an attitude.)

A similar bifurcation affects cases where individuals are said to *know* normative propositions. Suppose Jim and Brian both disapprove of capital punishment because they both think that it will lead to the deaths of innocent people. Then the following would seem a reasonable thing for Jim to say:

PUNISHMENT

Brian knows that capital punishment is wrong.

This is a case of mind-to-world fit. Disapproval by both individuals is based on a descriptive property which both think capital punishment possesses, that of leading to the deaths of innocent people. It thus seems reasonable to say that belief is involved, and by Jim’s lights, knowledge (note that Jim’s statement is a normative one – part of what is involved in his saying that Brian knows capital punishment is wrong is his expressing his own disapproval of capital punishment).

But now suppose that Jim and Brian’s negative attitudes to capital punishment are not based on any particular reasons, but are fundamental. Then, in the strictest sense, PUNISHMENT no longer seems like a reasonable thing for Jim to say, since, in this sense, Brian does not even have a relevant belief, let alone knowledge. And if belief is not involved, we may even feel that words like ‘true’ and ‘false’ are not, strictly speaking, applicable. But we can allow that by Jim’s lights, Brian does have knowledge in a looser sense of the term, and that in a correspondingly loose sense of the word ‘true’ - the minimalist one - he can speak of Brian’s belief as being true.

How important are the above considerations? Less so, I think, than might at first be supposed. For our discomfort with the application of the terms ‘true’ or ‘false’ to normative statements which, in their stand-alone uses, express attitudes fundamental to the speaker or to think that the propositions corresponding to them cannot be believed or known in the strictest sense does not seem to affect our *use* of them in discourse, which generally seems to proceed as it would in the case of descriptive statements. It is true that we do sometimes need to know whether someone’s attitude is or is not fundamental. But when this is so, the discussion is best framed precisely in terms of what attitudes people have and their reasons, or lack of reasons, for them, rather than by using normative statements. (Remember that, given the Redundancy Thesis, this is always an option for us.) Insofar as we think it right to express our attitudes using normative statements, the inferentialist model seems to capture correctly the manner in which we proceed.

But now consider the following argument:

HOPING

Jill hopes the outcome will be good.

The outcome will be good.

Therefore, at least one of Jill’s hopes will be fulfilled.

Now suppose that, in the present context, Jill thinks of a good outcome as one in which she will be happy, but we, the people making this argument, think of it as one in which *we* will be happy. Then the argument seems to fail, since we will only think its second premise true if we are going to be happy, and this is not what Jill thinks makes for a good outcome. The same sort of problem can be recreated for other propositional attitude verbs, such as ‘wants’, which represent some kind of desire for a state of affairs, where that state of affairs is normative. It arises because the normativity involved in a statement ascribing such a propositional attitude relates not to our own attitudes but to those of the subject of the propositional attitude. Let’s call this anomaly the *attitudinal shift anomaly* (ASA). Because of the ASA, arguments like HOPING are only valid in a context in which we can assume that our own attitudes and those of the subject of the propositional attitude verb are determined by the same descriptive (factual) considerations.

But I would argue that the ASA still does not undermine the inferentialist approach to normative statements. For it represents a *breakdown* in normal discourse involving such statements. As such, it does not reveal a need for a different way of conceiving of such discourse. But also note that, as in the previous case, the best course *in practice* for situations in which the ASA might arise is to shun normative statements in favour of descriptive statements about people’s attitudes and their differing reasons for them. This would remove the potential for confusion.

**3.4 Is normative truth relative?**

It is often maintained that moral truth, or indeed normative truth generally, is relative to such things as culture, society, individual or times. There are interesting and uninteresting relativistic theses. If someone were to suggest that normative truth was relative to an individual and a time merely because different people can have opposing attitudes to the same thing or because the same person can have opposing attitudes to the same thing at different times, this would be an uninteresting relativistic thesis. For it would be like saying that descriptive truth was relative to an individual and a time merely because different people can predicate incompatible properties of the same thing at different times or because the same person can predicate incompatible properties of the same thing at different times. Differences in attitudes do not in themselves create an interesting relativism, any more than differences in beliefs. However, on the expressivist view of norms defended here, there is an interesting relativistic theory about normative truth which, it seems at first sight, might be defensible. It is based on the fact that different individuals, or the same individual at different times, can have different fundamental criteria underlying their attitudes and hence their choice of which normative statements to accept. For example, one individual at one time may take the fact that a certain action was an instance of capital punishment as a conclusive reason for having a negative attitude towards it, while the same individual at a different time, or a different individual at the same time or at different times, may not, and this divergence may be fundamental in the sense that it is not due to any descriptive beliefs which the individuals concerned may hold about capital punishment. For this reason, it may seem that the truth or falsity of a normative proposition or statement is relative to a person and a time, as it is determined at least partly by the person’s fundamental attitudes at the given time, and there is no clear sense in which any one set of fundamental attitudes is, in absolute terms, better than any other. I could of course say that one was correct and the other incorrect as a way of agreeing with the first and disagreeing with the second, but in doing so, I would only be expressing my own attitude with regard to the matter. The relativity would still be there, but now with myself and the present moment occupying the role of the person and time involved, respectively.

There is more than one way in which such a theory might be developed. One of these has been called by Max Kölbel *indexical relativism* (Kölbel 2007). According to this view, normative statements are true or false in a way that resembles sentences containing indexical expressions. A token of the statement ‘I am feeling hot now’ produced by someone at a time when they are feeling hot is true, while a different token of the same statement produced by someone at a time when they are not feeling hot is false. In an analogous way, it might be said that a token of the statement ‘Capital punishment is always wrong’ produced by someone at a time when their fundamental attitudes require them to disapprove of all acts of capital punishment is true, while a different token of the same statement produced by someone at a time when their fundamental attitudes do not require this is false.

However, indexical relativism is very unattractive. For it would tell us that someone who says ‘Racism is a good thing’ speaks truly if their current fundamental attitudes require racism. This is not something we would want to say. It does not seem to be adequately mitigated by the fact that if we say ‘Racism is bad’ then we also speak truly if *our* current fundamental attitudes require being against racism.

However, there is a different way of developing relativism about norms, what Kölbel calls *genuine* relativism. According to this view, normative statement *types* are true or false depending on the fundamental attitudes, not of whoever says them on a given occasion, but of whoever is assessing them for truth or falsity. A given normative statement-type ‘*p*’ is true for *i* at *t* just in case the attitude(s) expressed by ‘*p*’ is/are required by the fundamental attitudes of *i* at *t*. On this view, we do not have to say that someone who says ‘Racism is good’ speaks truly. For we will only say this statement-type is true if it is true for us now.

However, I do not think this ultimately works either. To see why, note the following:

1. The relativity of the normative truth of statements would also produce a relativity of the truth of normative propositions, that is, not only ‘*p*’, but also *p*, is true relative to an individual at a time.
2. For any *p*, ‘*p’* is logically equivalent to ‘*p* is true’, which, on the view proposed, is logically equivalent to ‘*p* is true for me now’.

But it follows from (1) and (2) that for every normative proposition *p*, ‘*p*’ is equivalent to ‘*p* is true for me now’. But this makes the theory vulnerable to difficulties which parallel those arising for speaker subjectivism. (See Section 1.1).. For example, if I assert ‘Killing is always wrong’ and you assert ‘Killing is not always wrong’, then, according to the theory, my statement is equivalent to the claim that my fundamental attitudes always require disapproval of killing, while your statement is equivalent to the claim that your fundamental attitudes do not always require disapproval of killing. But if this were so, our two statements would not contradict one another, which they clearly do.

So, ultimately, I believe that even this more plausible relativistic account should be rejected. On the other hand, although my expressivism, for these reasons, should not technically be viewed as a form of moral relativism, it is a theory of the same general character. For it sees normative statements as expressions of subjective attitudes, and does not recognise any way of ‘adjudicating’ between conflicting attitudes that could be regarded as absolutely objective.[[36]](#footnote-36)

**3.5 Parfit on normative truth in expressivism**

Derek Parfit has raised issues connected with truth and falsity to argue that expressivism is less plausible for certain kinds of statements than it is for others. He considers two contrasting cases. In the first:

you and I are both non-fanatical supporters of the English football team. I say ‘Nothing good happened today.’ You say: ‘Not so. You’re forgetting that England won its game against Spain, so something good did happen today.’ I reply ‘That’s true.’ (Parfit 2011-2017, Vol. 3, 171)

Parfit accepts that in replying ‘That’s true’, he would be echoing my remark and thereby expressing his positive feelings about the result of the game. So in this case, he thinks the expressivist account is right. He also adds that if a Spanish person were to say that it was not good news, but bad news, he would not think this person’s claim was false, suggesting that we are not really dealing here with matters of fact. But he thinks this account fails for the following kind of case (or at least that it presents a challenge for the expressivist):

We learn that, after some shipwreck, several people have been rescued from the icy sea. (ibid. 172)

Parfit says that if he claimed that this was good news ‘I would not be merely expressing some favourable attitude. I would mean that we all had reasons to be glad that these people’s lives had been saved. If you … called my claim true, you would be implying that, if anyone claimed that this news was not good, that claim would be false’ (idem).

Contrary to Parfit, I believe the expressivist account applies in both cases with equal plausibility. In both cases, that is, to say that the news is good is to express one’s positive attitude to the news. Furthermore, with respect to both instances of good news, the following argument can be validly made:

GOOD NEWS

The news is good.

*Therefore,*  the statement ‘The news is bad’ is false.

The argument is valid because, as we saw in Section 3.3, given inferentialism, the word ‘false’ is used in broadly the same way when applied to normative statements as it is when applied to descriptive statements. But then what of Parfit’s claim that in the first example we would not describe as false the Spaniard’s claim that the news of England’s win is bad? I think that what we should say about this is that logic plus our own attitudes compel us to call the Spaniard’s claim false, but that it would be inappropriate to *accuse* the Spaniard of making a false statement. This is because we do not think that they have made any sort of error. For it was intrinsic to the example that we are not fanatics about supporting England, so we do not think that there is anything wrong with *other* people not supporting England. In fact, when discussing the matter with the Spaniard, as opposed to other England supporters, we are likely not to insist that the news was good. Instead, we are likely to revert to descriptive (i.e., non-normative) statements such as ‘From my standpoint, it is good news, but not from yours’.

Where descriptive statements are concerned, a person who believes a false statement is always, to that extent, in error, and this may explain why it feels uncomfortable to describe the Spaniard’s statement as false, since we do not think that they are in error. Of course, sometimes we think that people who believe normative propositions we would describe as false *are* in error. This applies to Parfit’s second case involving the people rescued from the sea, where we think that anyone describing the news as ‘not good’ would be in error, since we would disapprove of the attitude causing them to make that judgement.

Parfit also questions the ability of expressivism to account for the fact that normative statements can be mutually inconsistent. He criticises Blackburn for drawing an analogy between goals which are incompatible, and therefore jointly unrealisable, with inconsistent normative beliefs. Blackburn thinks the goals are bad in a similar way to that in which the beliefs are bad (Blackburn 1993, 189-190). Parfit rejects this (Parfit opus citus, 175). He notes that if I cannot save both my children from some terrible fate, and I naturally want to save both, the fact that these desires cannot both be fulfilled does not make either of them bad desires to have. In contrast, at least one of two inconsistent beliefs is always false, and hence bad in that way.

Now it is true that we would probably be reluctant to say that either of the desires in the example is a bad desire, since that might be thought to imply criticism of the agent, which would surely be unwarranted. On the other hand, it seems there is always a sense in which it *is* bad to have jointly unrealisable desires. It is bad for the person concerned, since by definition it entails that at least one of their desires cannot be fulfilled. Now the key claim of expressivism is best put by saying that normative statements express attitudes, rather than desires in the strictest sense. If someone in the present day condemns an action done by Julius Caesar, it does not follow that they have any *desires* relating to Julius Caesar. As explained in Section 2.3., the conative aspect of normativity may be much more indirect: in this case, for the condemnatory words to be correctly uttered, the speaker must think that Caesar’s action exemplifies a *type* of action which they would in general seek to avoid and advise others against. But then if the speaker says that Caesar should have done one thing, but also something else incompatible with it, they in effect commit themselves to an unrealisable pattern of action (under certain hypothetical circumstances at least) and to giving conflicting advice. Blackburn’s analogy therefore seems reasonable if it is recast in terms of conflicting patterns of (at least) hypothetical action rather than goals in the normal sense of that word. But it is important to note that in viewing the matter thus, the expressivist is not precluded from saying that having two inconsistent normative beliefs is a bad thing *also* because at least one of them must be false. For, as we have seen in this chapter (and as Blackburn would also agree), treating normative statements as expressions of attitudes does not rule out describing them as true or false (subject to the caveats made in Section 3.3).

**CHAPTER FOUR**

**Morality: structure and content**

**4.1 Moral statements**

Up to now, we have not talked very explicitly about morality. This needs to be remedied, since some might take the view that while norms in general are not absolutely objective, moral norms are. We must investigate whether the best account we can give of moral language really supports such a view.

Some of the examples of normative statements which we have used would naturally be taken to be moral statements. One is:

 PUNISHMENT

 Capital punishment is wrong.

However, ‘wrong’ here was not intended to relate specifically to moral wrongness, but to serve as a more general epithet – indeed the most general negative normative epithet that can be applied to an action, plan, policy etc. This is quite reasonable, since, though it is more than likely that someone who asserts PUNISHMENT does so from a moral perspective, this is not necessarily the case. For example, someone might coherently say:

NON-MORAL

Capital punishment is wrong. I don’t mean that it is *morally* wrong. I think it is wrong because, if capital punishment is available as a possible sentence, juries will often fail to convict when the evidence requires it, and this will lead to inadequate enforcement of the law, which would be contrary to our interests.

And of course there are many examples of using the word ‘wrong’ where no-one would even suspect a moral intent, such as:

 SPELLING

 His spelling was wrong.

However, in this chapter, our concern *is* specifically with morality. Note that if we wish to indicate that we are making a moral judgment using a general normative word such as ‘wrong’, we can do this simply by using the word ‘morally’ or some similar expression, as in:

 MORAL

 Capital punishment is morally wrong.

Or, alternatively, we may simply rely on the context (e.g., a moral philosophy seminar or a sermon) to indicate moral intent. Besides the word ‘wrong’, there are of course a range of more specific normative words which always or often carry a moral connotation, such as ‘wicked’, ‘evil’, ‘objectionable’ or ‘commendable’. Such words are not exclusively applied to actions: they can also describe things of other categories such as feelings or people. Again, in the case of a word like ‘commendable’, context indicates whether it is really morality which is being invoked or some other standard of appraisal.

This chapter is concerned with what I am calling the *intrinsic* features of morality. By this I mean those aspects of morality which are a direct consequence of the meaning of the word ‘moral’ and its synonyms. These features can be broadly divided into ‘substantive’ and ‘structural’.[[37]](#footnote-37) The former category concerns content: it is about what particular sorts of things we call morally good or bad, based on the meaning of the word ‘moral’. The latter has to do with general features of our concept of morality not related to content. Let’s consider these first. The following is not meant to be an exhaustive list of the structural features which have been, by some philosophers at least, considered essential to the concept of morality, but does contain arguably the most important ones. They mainly relate to the core notion of moral *obligation*.[[38]](#footnote-38)

Strong connections to reactive attitudes, such as approbation and blame. By default, if you are believed to have done something morally wrong, you are thought to deserve blame, and if you are believed to have done something morally good, you are often thought to deserve praise. Closely related to the former are strong connections to feelings of guilt and shame when we feel ourselves to have fallen short of our moral standards (these are much like attitudes of blame, but directed at ourselves).

The idea that moral obligations are ‘impersonal’ or ‘universalisable’: if I am morally obliged to act in a certain way, then anyone in a similar situation is morally obliged to act in the same way.

An overriding commitment to obligations viewed as moral: a moral obligation can only be trumped by another moral obligation, never by any other kind of consideration.

The idea that there is no such thing as ‘moral luck’. The thought here is that, for example, necessarily, if two people attempt a murder and, owing entirely to chance circumstances, one happens to be successful, while the other is not, this fact cannot make the former more morally culpable than the latter.

The idea that there are no moral conflicts. In other words, in any given situation, there is always at least one possible action which it would not be morally wrong to do (though there may be more than one, in which case, the agent is morally permitted to do any of them).

The idea that moral values have a particular kind of centrality in the life of an individual, in that they are the values which the individual would regret losing.

In the next section, I examine each of these features in more detail.

**4.2 Structural connections**

*Strong connections to reactive attitudes*

The close connection between moral language and reactive attitudes such as approbation and blame, as well as feelings of guilt and shame, can scarcely be denied. This is reflected in the availability of some very emotive words for expressing moral attitudes, such as ‘wicked’, ‘unconscionable’ and ‘egregious’. But even the emotive power of the word ‘moral’ itself is such as to make it likely that blame or punishment is a desired outcome for whoever has violated the relevant norm or norms. It is true that the presumption of blameworthiness is defeasible. It can be defeated, for example, if the person concerned was excusably ignorant of what their action involved or of its likely effects. It should also be noted that recognition of this conceptual connection does not settle the first order question of whether in fact blame or punishment are generally warranted for immoral behaviour. Indeed, a philosopher might deny, perhaps on deterministic grounds, that these reactions are ever warranted, but if they continue to describe the relevant behaviour as immoral, they will need to admit that they are not using that word in precisely its ordinary sense.

*Universalisation*

We now turn to the suggestion that moral claims are universalizable, so that to be morally acceptable, an action must be one we can approve of irrespective of the specific identity of the agent and of those affected by it. Originating in the work of Kant, this alleged feature of morality has been considered crucial by a number of philosophers. In fact some have thought it even has implications for the *content* of morality. For example, R.M. Hare uses the notion to ground a form of preference utilitarianism, according to which an action is morally right only if there is no alternative action which would promote better the preference-satisfaction of all those involved (Hare (1981), 107-111).[[39]](#footnote-39) Let us consider Hare’s argument for this.

To simplify his case, Hare develops it for a scenario in which only two people are involved, and then invites us to conclude that its extension to a multilateral scenario would be reasonable. In the two-person scenario, I am considering doing something *x* to another person (call him Joachim) which causes him to suffer. According to Hare’s universalisation condition, the moral acceptability of *x* cannot depend on its being Joachim in particular who would undergo the suffering and me in particular who would be the agent of it. It must be equally morally acceptable were our roles to be swapped. Therefore, in assessing whether it is right for me to treat Joachim in this way, I must imagine myself in Joachim’s place. Furthermore, I must do this in a way which accurately represents the impact of *x* on Joachim. (Let’s call this process *role reversal.*) Now Hare has a very strong claim about role reversal which is crucial for driving his argument. It is that anyone who manages to do it successfully will acquire the desires of the person he imagines being concerning how that person is to be treated, but applied to himself instead. (Call this the *empathy claim*.) In this case, Joachim’s suffering will make him not want *x* to be done to him, and so in role reversal, I will also want it not to be the case (and to the same degree as Joachim in the actual case) that I have *x* done to me in the counterfactual case where I occupy Joachim’s role. This puts me in a situation where I have two conflicted desires: my original desire to do *x* to Joachim and my new desire – created by role reversal – that *x* not be done to me in the counterfactual case. Hare maintains that I should deal with this conflict in the same way that I would any other intrapersonal conflict of two desires, which is to base my decision on the relative strength of the desires. This appears to yield the result that I may do *x* to Joachim only if my desire to do so is no weaker than his desire that *x* not be done to him. Generalising, this appears to lead to preference utilitarianism, according to which an action is morally acceptable just in case it leads to no less a degree of total satisfaction of the preferences of the people involved than any alternative action.

There are at least two points in Hare’s argument which seem vulnerable, threatening his case for preference utilitarianism. One is the empathy claim. Hare defends this claim partly by eliciting intuitive reactions to a concrete case (ibid. 94), and partly through the claim that ‘by calling some person “I”, I express at least a considerably greater concern for the satisfaction of his preferences than for those of people whom I do not so designate’ (ibid. 98). Whatever the merits of this latter claim, it is surely clear that neither it nor the intuitions about concrete cases, establish a feature of the empathy claim which is crucial to the success of Hare’s argument, namely, that the desire I acquire by imagining myself in Joachim’s place not to suffer the effects of *x* should be *as strong* as Joachim’s actual desire. This is implausible, since even if Joachim suffers quite badly as a result of *x*, then it is unlikely that my desire that it not happen to me is necessarily going to be as strong as his desire *while this suffering is happening* that it not be happening to him. For if that were the case, how, for example, could martyrs have chosen to let themselves be burnt at the stake instead of escaping that fate by confessing? (I assume here that the suffering involved of being burnt alive is so terrible that anyone with the opportunity to end it would do so.[[40]](#footnote-40)) We might wonder how such non-avoidance of extreme suffering is psychologically possible. Presumably, it has to do with the ability that all of us have to some degree to prevent ourselves from dwelling on unpleasant truths about the future. It is important to note that in such cases we can know what is going to happen to us at moments when we are not *representing* it to ourselves in imagination. Perhaps it will be insisted that we must have imagined it at least once, but the knowledge can outlive this act of imagination, so that even if the latter process generates a strong desire for avoidance, it may well be weakened subsequently.

Another problem with Hare’s argument is that even if the empathy claim is accepted, it does not lead to preference utilitarianism. Suppose that in Hare’s two-person scenario, my desire to do *x* to Joachim is in fact stronger than Joachim’s desire that *x* not be done to him. Then on the preference utilitarian view it is morally acceptable for me to do *x* to him. But this is not a scenario in which universality is satisfied. For it is still true that if I were in Joachim’s place, I would not want *x* to be done to me, and therefore (given the empathy claim) I do not want *x* to be done to Joachim in the hypothetical case where Joachim is me. Universality then fails, since it is not the case that I can approve of the action irrespective of the identity of the person who suffers as a result of it. The introduction of the procedure wherein I weigh my desire to do *x* in the actual case with my desire that *x* not be done in the case where I am Joachim is unmotivated. It would only be appropriate if I were to occupy, at different times, the position I actually occupy *and* the position occupied by Joachim.

But what of the principle of universality itself? Is it the case, as Hare and a number of other philosophers have maintained, that to be morally acceptable, an action must be one we can approve of irrespective of the specific identity of the agents and of those affected by it? Another, more formal, way of expressing the question is this: must it be the case that fundamental moral principles should not contain any irreplaceable uses of names or indexicals?

We do in fact see *generality*, if not universality, illustrated at least implicitly in any normative, or at least deontic, judgement. This is because, in a certain sense, moral judgements cannot be *gratuitous*. It is logically impossible to think an individual action wrong without its having some set of features such that one would also think another action having exactly the same set of features is wrong. (Note that the features involved could be very specific.) But the wrongness involved need not be morally based, and hence generality is simply a consequence of the nature of deontic normativity. It cannot therefore be the distinguishing mark of the moral. But universality, marked by the absence of irreplaceable names or indexicals, is obviously something more than generality. Is it essential to any moral principle?

Consider the following claim: ‘These moral rules apply in our society, but not necessarily in other societies’. It is very plausible to think that members of some societies have actually thought something like this. It is especially probable in the case of isolated tribes, who may have been surprised to discover that more technologically advanced societies even existed, and may have been quite unsure if ‘normal’ rules applied to them.[[41]](#footnote-41) It might be argued that this isn’t a genuine case of non-universality, as this frame of mind is compatible with the universal principle which says that a tribe’s moral rules bind its own members, but not the members of other tribes. But it is very implausible to suppose that they must adhere to any such principle, as they may have no views about what other tribes are or are not obligated to do. If so, their fundamental moral rules contain, at least implicitly, a reference to their own tribe. But how could they identify their own tribe, if asked to do so? Surely they would have to use either names or indexicals of some sort, referring either to themselves or perhaps the locality in which they live. This would be a clear violation of universality.

Another type of plausible counter-example to universality is the special consideration we often give to loved ones in our decisions about what to do. To invoke an example discussed by Williams, if I have to decide whether to save my wife or a stranger from a burning building, I will save my wife (Williams 1981, 18). It could be said that I adhere to a principle which goes something like ‘When you can, save others from danger, but prioritise the following people …’, where the ellipsis stands for a list of names including that of my wife and other loved ones. Could this be a moral principle? One’s first reaction is to say not, for it would only work for me – another person would have to write in a different set of names. But does everyone have to have precisely the same moral principles? There does not seem to be any conceptual fault in supposing that moral principles are personal or individual in this way. It would not, after all, prevent them from having some common core – in this case, the requirement to save people in general. It might be argued that there is a better principle, which *is* absolutely universal, namely, ‘When you can, save others from danger, but prioritise your loved ones’. ‘Your’ does not refer here to any particular person, but to whoever is using the principle. It is thus not an indexical, and therefore this principle could be claimed to be universal. What seems to make it superior to the earlier list-based one is that, unlike the latter, it encapsulates the reason *why* I should prioritise the particular individuals whom I do prioritise--because they are my loved ones. However, this advantage is also shared with the following principle which, like the list-based one, is not universal: ‘When you can, save others from danger, but prioritise the loved ones of Howard Simmons’. I can follow this principle, while you (for example) can follow a similar principle, but with your name substituted for mine. We could each be content with our respective principles, since for each of us, it explains why we favour the specific individuals that we do. Of course, in adhering to our respective principles, we would not be affirming the idea that we *all* ought to favour our loved ones (I, for example, would not be affirming that you ought to favour your loved ones), but why should this be thought necessary? At any rate, it is not clear that the resulting principles do not count as moral principles. More precisely, it is not clear that to describe them this way would be to misuse the word ‘moral’.

Williams himself argues that in justifying one’s decision to save one’s wife instead of a stranger, it is enough simply to say ‘She’s my wife’. To appeal to a more impartial principle is to have ‘one thought too many’ (idem). This is consistent with what I have said, except that we need to remember that, though a completely normal response, it is (like every other judgment) wholly subjective and someone who thought there *was* a need for further justification would not, in any absolute sense, be mistaken. It should also be acknowledged that Williams’ position regarding the word ‘moral’ is different from mine. Williams takes it for granted that morality is meant to be impartial by definition and hence concludes that morality is not overriding in this case. It is part of his attack on ‘the peculiar institution of morality’. I, in contrast, see nothing in the ordinary use of moral words (as opposed to the frequent pronouncements of moral philosophers) which compels absolute impartiality or universality in every case, and therefore do not see the stance of the person who favours his wife over strangers as being necessarily contrary to morality. However, this difference is perhaps not very important. Both Williams and I are opposing the hegemony of a certain kind of approach to making decisions – what label we use to describe this opposition is a relatively minor matter.

*Overridingness*

This property has already been mentioned in the previous section and, in fact, there is not an enormous amount to add. That moral obligations can never be overridden by non-moral considerations, such as prudence, is questionable, as a statement such as ‘I know that would be the *morally* right thing to do, but I don’t think it’s what you *ought* to do’ is surely intelligible. The thinking here might be that while morality is certainly one aspect of our lives (albeit perhaps a very important one), we are sometimes entitled to set it aside, especially when important needs of our own or of those close to us are at stake. Can a philosopher plausibly claim, on the contrary, that this violates the very meaning of the word ‘moral’, that overridingness should be seen as built into the concept of morality? Such a philosopher would then have to either dismiss the considerations alleged to override conventional morality, maintaining, for example, that it is never right to put our own needs, or those of our loved ones above the impersonal requirements of morality, or else admit that such considerations *can* be overriding and, therefore, that morality is not entirely impersonal. The first course is a hard one to sustain, since, as we have already seen, most people do think they have special obligations to their loved ones, or even to themselves, which, at least in extreme cases, override the more impartial demands of morality, and it is not clear what the defender of overriding impartiality could say to counter this.[[42]](#footnote-42) The second alternative seems better, and in its recognition that morality need not always be absolutely impartial, echoes a suggestion made in the last section.

*‘No moral luck’*

This is the idea that factors outside the agent’s control cannot affect how morally wrong or blameworthy an action is. The example given above was of two people who attempt murder, and one happens to be successful, while the other is not. The difference might be due to some random factor (e.g., an accidentally misdirected gun). The claim is that, *ceteris paribus*, the two attempted murderers have acted in an equally morally wrong manner and are therefore equally blameworthy. I sympathise with this view. but others feel differently, Bernard Williams being one (Williams 1981, 23). It is certainly not plausible to suggest that the idea is an essential constituent of the concept of morality.

*‘No moral conflicts’*

This feature has perhaps the loosest connection of all with the concept of morality, since most speakers (I suggest) would not find it at all counter-intuitive to think that there could be an irresolvable moral dilemma.

People often seem to find themselves in painful moral conflicts. For example, I may not be able to tell the truth without seriously offending someone who does not deserve to be offended: should I tell the truth or not? We often feel in such situations that there must be a right answer, if only we had the moral insight to discover what it was. (Of course, there may be people who do not think of this example as a moral conflict: they may feel sure they know what to do; but it is enough for the purposes of the example that *some* would experience it as a genuine dilemma.) If this feeling that there must be a right answer is correct then, *objectively*, there is no moral conflict, however conflicted we may be in our own minds when we consider the matter. It should be clear, however, that this view is incompatible with the meta-ethical theory defended in this book. If there is no answer to any normative question which is correct in an absolutely objective sense, then *a fortiori* there is no such answer in situations where we feel morally conflicted.

*Centrality*

This is how I refer to Blackburn’s idea that our moral values are a proper subset of our desires, those which we would like to preserve (Blackburn 1998, 67). For most of those who are kind to others, for example, this characteristic is one which they would be unwilling to lose, unlike a desire for ice-cream, say, which many would not regret losing. But here too we have a feature whose conceptual connection with morality is not undeniable, since a person could intelligibly and sincerely say ‘I wish I weren’t such an honest person – it makes my life very hard’. Indeed, some people do say or think such things, and it would be highly implausible to suggest that in doing so, they are excluding honesty from the category of moral virtues.

**4.3 Substantive features**

Our examination of the possible structural components of the concept of morality has been overall somewhat disappointing. In most cases, the findings have been rather minimal. I turn then to those aspects which relate to content. This has to do mainly with what sorts of actions are morally right or wrong, but also with what sorts of desires or character it is morally good or bad to have. Are there any general claims we can make about this just by examining what we mean by the word ‘moral’? For some reason, philosophers have tended to take less interest in this subject than in the structural features (especially universalisation). I don’t mean, of course, that they have not been interested in the content of morality. Indeed, many have devoted a great deal of attention to it. I mean that when it is a matter of trying to discern what can be said about morality on the basis of the *meaning* of the term ‘moral’, structural matters have achieved much more prominence than matters relating to content.

It does seem to be the case, however, at least that there are certain content-related features which are such that, if a set of norms lacks them, we are very unlikely to call it ‘moral’. For example, if a set of norms says nothing about how we should treat other people in high-stakes situations, we will not call it moral (etiquette, for example, is not a form of morality). If a set of norms allows indiscriminate killing or cruelty, we will not call it a moral code (the word ‘indiscriminate’ is obviously important here – most moral codes allow killing and injuring people under certain specific kinds of circumstances). There are also aspects of the nature of moral *virtue* which place constraints on how we use the word ‘moral’. As Philippa Foot has pointed out, for some behavioural pattern to count as a virtue, it must be in general rather *hard* to practise (Foot 2002, 9-10). Courage, for example, requires one to overcome one’s fear. This is not to deny that there are individuals who find it somewhat easier to act virtuously than others (ibid., 10). For example, people with more empathy will find it easier to act considerately towards other people. It is rather that moral virtues tend to be such that their manifestation *normally* runs counter to people’s inclinations. By definition, being moral is not an easy thing *in general*.

The meaning of the word ‘moral’ seems to have changed over time. A century ago, the use of the word would have carried with it strong associations with rules of sexual conduct. This is less true now. Currently, the aspects of conduct which ‘moral’ most prominently bring to mind are those pertaining to kindness, compassion and honesty.

The word ‘moral’ does not have precise boundaries for its correct application, and therefore, despite the foregoing considerations, we cannot formulate precise necessary and sufficient conditions for something to be morally right based purely on the meaning of the phrase ‘morally right’ (unless we cheat by including words in the conditions which themselves have moral import). The use of the word ‘moral’, like the word ‘game’ in Wittgenstein’s classic account, is perhaps based on the recognition of ‘family resemblances’ between a wide range of different types of cases (Wittgenstein 1963, sec. 66-7, 31-2; Hare 1981, 54). Though we can often point to common features of different types of cases in which the word ‘moral’ is used, they do not hold with the sort of reliability that would enable them to form the basis of necessary and sufficient conditions.[[43]](#footnote-43)

Call the totality of types of attitudes that can legitimately be called moral MORALITY. In view of the foregoing considerations, MORALITY is a somewhat indeterminate thing, but it is not *wholly* indeterminate: it constitutes a loosely defined set of constraints along the lines we have been discussing. Roughly speaking, I propose that, in standard cases, we understand a speaker who says, for example, ‘It is morally wrong to steal’ as expressing the following:

1. a negative attitude to stealing;
2. a general willingness on the part of the speaker to respond to those who steal with the reactive attitude of blame (albeit subject to certain exceptions);
3. that a negative attitude to stealing belongs to MORALITY.[[44]](#footnote-44)

In other words, the assertability condition of this statement is that the speaker has each of the attitudes in clauses (1) and (2) and that a negative attitude to stealing belongs to MORALITY.

Parallel to the treatment of the general wrongness of individual actions in Section 2.3, we can say that a speaker who says, for example, ‘That action was wrong’ expresses that the action referred to belongs to some action-type *x* such that the assertability condition of a statement attributing moral wrongness to *x* is satisfied.

Another important moral notion is that of a ‘morally good person’. But given the foregoing analysis of thinking an action-type morally wrong, this is straightforward to handle, and was in fact already given in Section 2.2: a morally good person is someone who generally performs morally good actions and avoids morally bad ones.

**CHAPTER FIVE**

**Alternative Theories and their Shortcomings**

**5.1 Introduction**

In this chapter, I consider alternatives to expressivism for the understanding of normativity and explain why I believe expressivism is superior to all of them. It will not be possible, of course, to examine every alternative that has ever been suggested, but I will look at a good sample.

The theories are summarised in the table below. Each entry is either a specific theory or a group of similar theories.

|  |  |
| --- | --- |
| **Theory** | **Explanation** |
| Speaker subjectivism | The theory that a normative statement describes the speaker’s attitudes. |
| Basic naturalism | The theory that every normative/moral statement is logically equivalent to some naturalistic statement, i.e., a claim which could be established or refuted by empirical investigation. |
| Cornell realism | A form of naturalism which identifies moral properties with natural properties (properties that can be investigated by natural/empirical science), but without asserting logical equivalence. |
| Intuitionism | The theory that moral statements are descriptive, but not naturalistic, and that some moral statements, so understood, are true. |
| The error theory | The theory that moral statements are descriptive, but not naturalistic, but that *no* moral statements, so understood, are true. |
| Constructivism | The theory which, though not treating normative/moral statements as simply equivalent to naturalistic statements, nevertheless maintains that they are made true or false by facts about human choices or conventions. |
| Universal prescriptivism | A theory developed by R.M. Hare, according to which moral ‘ought’ statements are ‘universal’ prescriptions for action.  |
| Hybrid theories | Any theory which maintains that even what are normally regarded as thin normative statements (ascriptions of goodness or rightness, for example) have expressive or prescriptive *as well as* descriptive aspects to their meaning. |

As the above table makes clear, some of the theories tend to be offered as theories of normative statements in general, whereas others are put forward as theories of specifically moral statements. In the case of a few theories, both of these options are possible.

Most of the theories listed here will be discussed in this chapter. But not all. Speaker subjectivism has already been refuted in Chapter One. There will also be no discussion of Hare’s universal prescriptivism, as I have already raised doubts in Chapter Five about the reasonableness of accepting a universality constraint for moral statements. (However, it should be remembered that my brand of expressivism contains an element of prescriptivism in its reference to advice-giving in the analysis of thinking an action-type wrong.)

**5.2 Basic naturalism**

The first two theories to be examined are both forms of naturalism, that is, both try to understand normative or moral properties in terms of naturalistic properties open to investigation by the methods of empirical science. The difference is that basic naturalism presents this link as a logical or analytic equivalence, whereas Cornell realism treats it as an empirically discoverable fact. In this section, I look at basic naturalism.

Basic naturalism says that a normative term can be given a naturalistic *analysis.* Let us focus our attention on the most widely cited case, the word ‘good’. What sort of analysis might seem plausible for this word? Clearly something as crude as equating ‘*x* is good ‘ with, say, ‘*x* is desired’ won’t be adequate, as there is no contradiction in supposing both that *x* is desired (by someone or other) and that *x* is not good. A more sophisticated suggestion might be to replace ‘*x* is desired’ with something like ‘*x* is, or would be, desired by someone after full rational deliberation on the basis of all the relevant facts’.

But this proposal is vulnerable to a dilemma. How is the word ‘rational’ being used here? There are two broad possibilities: a normative and a naturalistic use. If it is being understood normatively, then the schema will not fulfil its purpose, as the whole point of the (basic) naturalist’s project is to translate normative statements into descriptive ones.[[45]](#footnote-45) Alternatively, ‘rational’ could be given a naturalistic reading, by specifying in an attitudinally neutral way the kinds of procedures which we would normally regard as constituting rationality. But then the naturalist’s proposal is vulnerable to G.E. Moore’s ‘open question’ argument. Moore argues that any proposed definition of ‘good’ which construes ‘*x* is good’ as being equivalent to ‘*X* has property *p´*, where *p* is some natural property, would be subject to the objection that one could always intelligibly ask ‘Are things which possess *p* good?’, in which case it could not be true that being ‘good’ was equivalent in meaning to possessing *p* (Moore 1903, Sec. 13). We can easily see that this objection applies to the definition of ‘good’ proposed above, when its use of the word ‘rational’ is interpreted naturalistically. For we could always intelligibly ask in that case whether something which is desired by someone after full rational deliberation on the basis of all the relevant facts was actually good. Thus the proposal fails both with a normative and a naturalistic interpretation of the word ‘rational’. And the same applies to any other proposed naturalistic analysis of the word ‘good’.

In contrast, expressivism manages to evade this argument altogether, since it does not even attempt to give an analysis of ‘good’ or any other normative word. Rather, it explains the meaning of such words by reference to the state of mind expressed by the person using them. On the other hand, it has been argued that expressivism does have a kind of open question problem. For even when it is known that someone has a positive attitude towards something, it may seem intelligible to ask if they really think it is good.[[46]](#footnote-46) But I think that when we examine particular cases of this, it is not so clear that expressivism is undermined. Suppose, for example, someone were to say:

Keji has a positive attitude to homework, but does she really think it is a good thing?

Why might they say this? Perhaps what they mean is that Keji is very enthusiastic when doing her homework, but, despite that, she may not think it actually benefits her. But then there is a discrepancy between the way in which Keji is said to have a positive attitude towards homework in the first clause and the kind of goodness-ascription involved in the second clause. The former is purely to do with enthusiasm in pursuing the activity and the latter with benefit. This suggests that the expressivist’s central claim about ‘good’ actually amounts to something like this:

Thinking something good in a certain way is having a positive attitude to it *in the same way*.[[47]](#footnote-47)

This formulation, I believe, survives any attempted undermining by the open question argument, since it would not be intelligible to say:

Keji has a positive attitude to homework in a certain way, but does she think it good in the same way?

Another difficulty for naturalists which expressivists do not face is the problem of accounting for the property of certain kinds of normative statements noted in Section 2.3 and referred to as ‘motivational load’. The normative statements in question are those which entail that under certain possible circumstances, the speaker ought to behave in a certain way.[[48]](#footnote-48) The clearest case of this would be a statement which directly asserts that the speaker has some obligation, such as ‘I must leave here now’. It cannot be correct for me to assert this statement unless I am strongly motivated to leave. Another, more indirect, type of case is a general statement which makes a judgement about an action to be carried out under circumstances which the speaker may find herself in at some time, e.g., ‘If someone makes a promise, she should keep it’. A person should not assert this unless she is at least somewhat motivated to keep her promises.

Some opponents of expressivism have doubted that the motivational load property is really an essential property of any normative statement. Following standard terminology, let’s refer to those who deny this as normative *externalists* and those who accept it (at least for the relevant proper subset of normative statements) as normative *internalists*. It might be thought that this is an issue which could be resolved empirically by undertaking a suitable study. And indeed such studies have been done. Caj Strandberg and Fredrik Björklund set up one in which subjects were asked to consider the following scenario:[[49]](#footnote-49)

Anna is watching a TV programme about a famine in Sudan. In the TV programme, it is shown how the starving are suffering and desperately looking for food.

At the same time, Anna is not motivated at all, not to any extent, to give any money to those who are starving.

The subjects, none of whom had ever been exposed to the controversy surrounding normative internalism, had to answer the following question: 'Could it be the case that Anna thinks she is morally required to give some of her money to the starving even if she not motivated at all to do so?'

The results that Strandberg and Björklund obtained appear to provide little support for internalism—indeed, quite the opposite. 76% answered 'yes' to the question. In other words, 76% were prepared to say that it could be the case that Anna thought herself morally required to give some of her money to the starving even if she was not motivated at all to do so.[[50]](#footnote-50) But does this really refute normative internalism?[[51]](#footnote-51) We could argue that those who answered ’Yes’ were simply not interpreting the phrase ’morally required’ in a normative way. They may simply have been reading it as concerning what Anna thinks is required by the moral code of her society. The results of the study therefore do not contradict the fact that genuinely normative statements express attitudes. This is all the normative internalist really needs. Normativity just is the use of a statement to express an attitude. (Remember from the first chapter that normativity is that thing which ‘That is a terrible joke’ and ‘Murder is wrong’ have in common which is lacked by, for example, ‘Water is a compound’ and ‘Mars has two moons’, and it is hard to think of what that thing might be other than the expression of an attitude.) But surely, having a negative attitude to oneself doing a certain type of thing or to everyone’s doing a certain type of thing analytically entails being motivated not to do that type of thing. So to the extent that normative statements carry motivational load, expressivists are able to explain this quite satisfactorily.

In contrast to expressivism, basic naturalism appears unable to explain the motivational load property of (some) normative statements. This is because it treats these statements as equivalent to certain descriptive statements and it seems that no purely descriptive statement has motivational load. An example of a naturalistic analysis, suggested above, for ‘*x* is good’ is ‘*x* is, or would be, desired by someone after full rational deliberation on the basis of all the relevant facts’. But someone could accept that *x* has the latter property without being motivated to act in any particular way in relation to *x*.[[52]](#footnote-52)

Despite these problems, basic naturalism still has its defenders. A recent example is Peter Railton (Railton 1986), who describes his theory as a version of ‘moral realism’. Railton first develops a naturalistic analysis of non-moral goodness for a given individual. He suggests that *x* is non-morally good for *i* just in case *x* would satisfy an ‘objective interest’ of *i*, where an individual’s objective interests are based on a ranking of their wants generated by a hypothetical version of that individual endowed with perfect cognition and all relevant information. Objective interests may be prima facie or they may be ‘on balance’, the latter providing a concept of ‘*the* good for *a*’ (Railton 2007, 192). Railton views this idea as generating a theory of norms based on the concept of instrumental rationality. He extends this notion to cover morality: ‘moral norms reflect a certain kind of rationality, rationality not from the point of view of any particular individual, but from what might be called a social point of view’ (idem).

Railton’s analysis of normative words is interesting and subtle – more so than is apparent from the above brief sketch. But is it really true to our actual use of these words? Railton claims, in effect, that claiming something is ‘part of one’s good’ carries motivational load:

Suppose that one desires *X*, but wonders whether *X* is really part of one’s good … If one were to learn that one would still want oneself to want *X* in the circumstances were one to view things with full information and rationality, this presumably would reduce the force of the original worry. By contrast, were one to learn that when fully informed and rational one would want oneself *not* to want *X* in the circumstances, this presumably would add force to it (idem).

Wants are, of course, intrinsically linked to motivation. Despite this, Railton’s concept of the good is not strong enough to ground motivational load in the sense explained earlier. This needs it to be logically *guaranteed* that the subject will be motivated towards something if he thinks it is a part of his good. And given Railton’s concept of a person’s good, this would only obtain if it was already built into the concept of rationality. But, as we remarked earlier in connection with a similar naturalistic analysis of goodness, this would defeat the object of naturalism, which is to define goodness and other normative concepts in non-normative terms.

Railton admits that his account is vulnerable to the open question gambit, since, for example, knowing that a practice would be rational from an impartial point of view still allows someone to intelligibly ask whether it would be right (ibid. 203). Since Railton grants this point, how does he sustain his commitment to this meta-ethical theory? One way is to point out that, although it is true thar not every rational agent will be motivated to do what is morally right according to his definition, the importance of that definition is guaranteed by the fact that we are at any rate *often* concerned with whether an action ‘is justifiable from a general rather than merely personal stand-point’ (ibid, 202). He also somewhat evades the force of the open question and motivational load arguments by stating that his definitions of non-moral goodness and moral rightness are in effect reforming definitions (ibid. 203). His point is perhaps that these concepts are based on the familiar concepts of goodness and rightness but involve differences from the latter which make them in some sense more fit for purpose. Precisely because they are not exactly our ordinary concepts of these notions, we might allow that they do not have to incorporate motivational loadedness, nor is it fair to oppose them with the open question argument. But there is a general argument which can be used against reforming definitions of any term, which is that in every case that I know of, the new definition has never been adopted by anyone other than the philosopher who has proposed it. Having said that, Railton’s emphasis on the empirically determinable factors that make lives go well or badly is, I think, very much to be welcomed. But I would argue that no reason has so far been given why we cannot simply adhere to our ordinary concept of moral rightness, explained according to the theory advanced in the previous chapter (itself based on the more general expressivist theory defended in Chapters Two and Three), while urging the acceptance of moral norms which put these empirical factors centre stage.[[53]](#footnote-53) Such urging involves taking what is in effect a substantive position in favour of the idea that we should resolve moral questions through appeal to such factors. However, we should not be too optimistic about how far this will take us. In a later chapter, we will observe how moral theory is beset with conflicts which do not appear to be susceptible to any resolution that can be regarded as objectively correct in an absolute sense, even supposing we do make empirical considerations paramount.

**5.3 Cornell realism**

Railton’s version of naturalism, despite its sophistication, is still a form of what I call ‘basic’ naturalism, as it offers a way of translating normative/moral statements into allegedly equivalent descriptive statements (although, in Railton’s case, as we have seen, this translation is ultimately considered as providing a reforming definition, rather than a characterisation of our actual use of normative statements). Because of the difficulties surrounding basic naturalism, a number of authors have developed a different approach to securing the naturalist’s goal. The key naturalist claim for such authors is that a given normative property is identical to a certain descriptive property, which, they argue, does not require supposing that a normative claim is identical in meaning to any descriptive claim. Such theorists often cite the fact that empirical science provides examples of many identities, such as that between lightning and a certain kind of electrical discharge, which are discovered *a posteriori* and are thus not predictable from the meaning of the words involved, assuming that by ‘meaning’, we are referring to what ordinary users of the term (not necessarily scientists) need to understand in order to be able to use it competently. Another example would be the treatment of natural kind terms. Water is identical to H2O, but someone could be said to know what ‘water’ means without knowing this. The suggestion is that the identity between moral goodness or rightness on the one hand and some descriptive property on the other is of the same general kind. Thus, for example, Richard N. Boyd argues that moral goodness should be identified with a ‘cluster of goods’,[[54]](#footnote-54) based on fundamental physical and psychological needs, together with ‘the homeostatic mechanisms which unify them’ (Boyd 2007, 175). Other forms of Cornell realism would typically identify goodness with some other property, but it would always be something which could be identified with goodness only *a posteriori* on the basis of empirical investigation.

Does Boyd’s theory succeed in overcoming the difficulties surrounding basic naturalism? I don’t think it does. For one thing, the open question argument is still a problem. People were obviously using the word ‘good’ intelligibly before the relevant cluster was identified. In describing something as good, they meant, according to the theory, something like ‘possessing one of the properties entailing the satisfaction of physical and psychological needs, whatever these may be’. But then the open question argument asserts itself, for it would be intelligible to say ‘*x* possesses one of the properties entailing the satisfaction of physical and psychological needs, but is *x* good?’. And this is a problem which applies, not only to Boyd’s version of Cornell realism, but also to every other, since such theories all have to provide some account of how the word ‘good’ is understood before the actual nature of goodness is identified.

In addition to this, the problem of motivational load still seems to be present, as the theory treats moral statements as objective statements of fact, rather than as the kind of thing (such as an expression of attitude or a prescription) that could carry a motivational load.

Furthermore, it is not clear that a version of Cornell realism is really a *theory* of anything. Obviously there are interconnected human needs. Finding out what these are (to the extent that we do not already know this) is important and will quite likely involve some non-obvious science. But it is not really clear how this is analogous to finding out that lightning is a kind of electrical discharge or that disease is the effect of micro-organisms. It seems to be an *application* of science, rather than a scientific theory *per se*.

A further point is that there are, as I shall argue in the next chapter, radical indeterminacies concerning what is and what is not good. Boyd might object that the existence of hard cases does not in itself undermine his theory. After all, there may be such cases arising for ordinary natural kinds as well. For example, it seems indeterminate whether heavy water is really *water*. And it is certainly true that (to put the matter in subjective terms) the existence of such cases in relation to goodness does not prevent us from favouring that which conforms to the suggested conception of moral goodness in straightforward cases, and perhaps basing our lives around it to the extent that we can. However, it must surely be acknowledged that many hard cases in ethics are considerably more weighty than the question of whether heavy water is really water. Furthermore, they are hard precisely *because* they are not resolvable by empirical methods.

**5.4 Intuitionism**

Another theory of moral statements is that, although they do not attribute properties which are to be understood empirically, they do make claims which are true or false in a more than minimal sense. This view, associated with authors such as G.E. Moore and W.D. Ross, is often referred to as *intuitionism*.

Although he does not use the term ‘intuitionism’ (he prefers ‘moral realism’), Russ Shafer-Landau is an example of a contemporary philosopher who defends this theory. One thing he does to defend this position is to address a couple of arguments often used against realist positions in ethics (that is, positions which construe moral statements as descriptive, though not necessarily empirical). These are the argument from moral disagreement and the argument from the causal inefficacy of moral facts.

The argument from moral disagreement appeals to the fact that there is widespread and persistent disagreement about morality. At first glance, this might seem to be untrue. After all, almost everyone agrees, for example, that it is morally wrong to kill someone or to break a promise. These prohibitions are felt to hold under most circumstances, though not all. But under *exactly* what circumstances is it morally permissible to break any of these moral rules? On that point there is likely to be much disagreement. Also, moral philosophers, as opposed to ordinary folk, disagree greatly about the most basic principles which appear to underlie, and perhaps justify, the more familiar rules. (The contenders include utilitarian, deontological, contractarian and virtue-ethical principles.) It is often argued that the existence of such widespread disagreement about both the content and the fundamental principles of morality renders a realist account of morality implausible. An appeal can made here to the case of science. If scientists regularly and radically disagreed about the facts regarding physical particles or the biological basis of life, we might begin to question whether their various postulated entities and processes and alleged associated properties weren’t just a figment of their imagination.

Shafer-Landau finds such considerations uncompelling (Shafer-Landau 2007, 214). The basis of his riposte is that the same argument could be made about philosophy generally, so that accepting the argument would imply taking an anti-realist view of philosophy in general, which he considers implausible. He also maintains that if I have a belief which is justifiably rejected by everyone else, I might still be justified in holding on to this belief. To make his case here, he presents an example in which I have a clear memory of an event, while others have good reasons to doubt it ever occurred – this, in itself, he argues, is not a good reason for me to doubt my memory, and hence my belief, that it did indeed happen (ibid., 216).

As for the alleged causal inefficacy of moral facts, Shafer-Landau argues that this cannot convincingly be used as a reason to doubt the reality of such facts. The argument, as stated by Shafer-Landau, is that if the existence of something is to be construed realistically, then it must have independent causal powers, and, since moral facts do not have independent causal powers, it follows that their existence (if they can be said to exist at all) cannot be construed realistically (ibid, 217). Shafer-Landau’s response to this argument is, in effect, that its first premise is far too strong to be at all plausible. Accepting it would mean that we would have to suppose that only the fundamental objects of physics are real; for the objects studied by the special sciences, such as chemistry and biology, do not have independent causal powers: their properties supervene on the properties of objects studied by physics. Additionally, it would mean that we would have to deny the reality of other kinds of norms in addition to moral ones – for example, the epistemic norm according to which we ought to accept some particular argument because it is valid, as this too is causally inefficacious. These conclusions are not at all plausible, according to Shafer-Landau (ibid., 218).

As far as the argument from widespread persistent disagreement is concerned, I am inclined to think Shafer-Landau is right: the existence of such disagreement in some area of discourse might superficially suggest that there is no (absolutely objective) right or wrong about it, but it is not a strong argument for the latter conclusion. The argument from causal inefficacy is harder to assess, as it would take us into areas of ontological reasoning which go far beyond the scope of this book. However, Shafer-Landau’s claim that epistemic norms must be objectively real even though they are not causally efficacious, is questionable. On the contrary, all norms, I would argue, should be construed in expressivist terms, and this includes epistemic norms.[[55]](#footnote-55)

So why do I reject intuitionism if I am not prepared to rely on either of the two arguments discussed by Shafer-Landau? This is because intuitionists do not provide a clear account of what moral normativity actually is. An expressivist account was given at the end of Chapter Four. Naturalists have an alternative conception appealing to empirical properties. But what is the intuitionist’s conception? It is not clear. Moore thought that ‘good’ was indefinable (Moore 1903, Sec. 6) and other moral words sould be defined in terms of it. But this leaves a big explanatory vacuum at the heart of intuitionism, rendering it, in my view unacceptable as a meta-theory of morality.

**5.5 The error theory**

The error theory combines the intuitionist’s claim that moral statements are descriptive, but non-naturalistic with the claim that they are always false. The first claim is one we have already objected to on the gounds that it leaves the nature of moral normativity unexplained. The second claim marely makes it more implausible still: it is very hard to believe that every moral statement that anyone has ever made is false.

In Section 1.3, I endorsed the ‘Redundancy Thesis’, which says that if we removed normative statements (a category which of course includes moral statements) from our language we would, in general, lose only stylistic variety, rhetorical force, elegance or succinctness. It may seem that the RT is a kind of error theory in that it entails, not that ordinary folk are in error when making moral statements, but that those who think there are objectively real moral norms are in error in thinking *that*. Another reason for thinking that my view is a kind of error theory derives from the fact, noticed in Section 3.5, that normative statements expressing fundamental attitudes do not express beliefs or knowledge, and are not true or false in quite the fullest sense, unlike normative statements expressing non-fundamental attitudes. This might be considered grounds for thinking that someone who makes a normative statement ‘*p*’ expressing a fundamental attitude has misspoken, since for any statement ‘*p*’, it is a logical truth that asserting ‘*p*’ commits one to accepting “’*p*’ is true”. As I made clear in Section 3.5, I do not think it is determinately the case that such statements cannot be described as true or false, but insofar as a case can be made for this, it would provide another reason for thinking that my theory is a non-standard kind of error theory (non-standard because it does not treat all moral statements as straightforwardly false).

**5.6 Constructivism**

Constructivist theories are theories of morality taking their inspiration from the fact that many norms can be explained by reference to human conventions. For example, the norms of etiquette and of baseball can be explained by reference to the respective conventions governing these areas of life (Shafer-Landau and Cuneo 2007, 80). The claim is that this also applies to morality. Shafer-Landau and Cuneo explain such theories in terms of the idea of a *conceiving-dependent* fact:

…something is a conceiving-dependent property just in case if something exemplifies that properly, then it does so in virtue of the intentional attitudes taken toward it (or that would be taken toward it) by some actual or idealised (human) agent(s) … a conceiving-dependent fact *that x is F* is one in which a thing, x, instantiates a conceiving-dependent property *F*. (idem)

Shafer-Landau and Cuneo then define moral constructivism as stating that ‘every moral fact is either itself a conceiving-dependent moral fact or explained by a conceiving-dependent moral fact …’ (idem). If conceiving-dependent moral facts are understood in terms of the intentional attitudes of actual human agents, one gets a theory like that of Gilbert Harman, which views morality as the result of an implicit bargain (ibid., 81, 84-92) or like that of Christine Koorsgaard, whose Kantian view sees morality as the necessary consequence of our being rational agents (ibid., 81, 93-106). An example of the alternative conception, according to which morality is the product of the intentional attitudes of *ideal* agents is the theory of Michael Smith, according to which an act is right for an agent *i* in circumstances *c* just in case and because an idealised counterpart of *i* would desire *i* to act in that fashion when in *c*, where idealised agents are conceived of as fully rational (ibid 82; Smith 1994).

Constructivism may seem an attractive meta-theory of morality. It seems to secure genuinely truth-evaluable moral facts, while at the same time validating the intuition that there would be no such thing as morality if there were no human agents (or at least rational agents resembling human beings in important respects). Morality is a human creation, according to this view. Also, constructivism is not metaphysically extravagant. It is not committed to the existence of a moral realm beyond the bounds of the empirically knowable, as intuitionism appears to be (ibid., 80).

Despite these advantages, constructivism does not score highly as a moral meta-theory. For one thing, It does not recognise motivational load. (To take Smith’s theory as an example, I could know that my idealised counterpart would want me to do some action, but nevertheless be completely unmotivated to do it.) It is also vulnerable to the open question argument. (Again, in reference to Smith’s theory, it is intelligible to ask ‘Is an act desired by my idealised counterpart the right act under the circumstances?’) These failings entail that it compares poorly to expressivism and prescriptivism, which avoid these pitfalls. Yet, as we have seen, theories of the latter kind can be developed in such a way as to secure a robust sense in which normative statements, including moral ones, can be treated as truth-evaluable, so constructivists cannot claim this as an advantage for their theory compared with expressivism and prescriptivism.

I believe that constructivism goes wrong because, although it has a theory of norms (as the products of certain sorts of actual or idealised human activity), it does not have a theory of the meaning of normative *statements*. The latter do not merely function to describe existing norms. They involve *committing* oneself to norms. To suppose otherwise is to fall victim to the error identified by Hume of thinking that an ‘ought’ can follow from an ‘is’, something which is not possible as long as the ‘ought’ is a genuinely normative (i.e., ‘committed’) one. (An example of a non-normative ‘ought’ is the ‘ought’ in ‘I know I ought to keep my promise, but I am not in any way motivated to do so’.)[[56]](#footnote-56)

**5.7 Hybrid theories**

*Hybrid theories* are meta-ethical theories which characterise normative (or moral) statements as expressing attitudes or making prescriptions, while at the same time maintaining that they are *also* descriptive. (Hybrid theorists will therefore reject my way of dividing statements into the normative and the descriptive, since on my conception, the essence of a descriptive statement is simply that it does not express any attitudes. Instead, a hybrid theorist can say that normative or moral statements are descriptive in the sense of having more than minimal truth conditions.) Hybrid theorists are impressed by the fact that normative statements seem to have some properties associated with descriptive statements (for example, truth-aptness, and the ability to be embedded into longer statements), but also properties associated with non-descriptive statements, performing such functions as expressing attitudes or giving prescriptions (for example, motivational load). They believe that the best explanation for this is that normative statements belong to both camps.

A number of different hybrid theories have been proposed, and it is not possible to cover them all here. I shall briefly look at three prominent ones.

First, a very early one: that of Paul Edwards in his book *The Logic of Moral Discourse* (Edwards 1955). I choose to mention this theory because it recognises a very important distinction which is not recognised by many theorists, but which is virtually identical to one which I drew attention to in Section 3.3. This is the distinction between fundamental and non-fundamental moral judgements (ibid., 182-192). An example of the former would be ‘Happiness is intrinsically good’, while an example of the latter would be ‘She is a good person’. Non-fundamental judgments are, whereas fundamental judgements are not, made on the basis of reasons. Edwards maintains that while both sorts of judgements have emotive meaning in the sense that they serve to express emotions or attitudes, fundamental judgments have emotive meaning *only*. In contrast, non-fundamental judgements, are, as a matter of meaning, what Edwards calls ‘objective’, or in other words, they have descriptive truth-conditions: there are descriptive criteria by which we can determine, for example, whether someone is a good person. Thus fundamental moral judgements have a different sort of meaning from non-fundamental ones.

On this last point I disagree with Edwards. To my mind, the meaning of all normative statements (including moral statements) lies in the fact that they express attitudes. Suppose Pete says ‘She is a good person’ on the basis of very different descriptive criteria for morally evaluating people from those employed by Edwards and me. For example, suppose Pete thinks that being bigoted is a good-making quality. In such a case, I would not say that Pete has misused the expression ‘good person’. The objection to Pete is not his use of language, but his attitudes. Because of this I prefer not to explain the different relationships which fundamental and non-fundamental statements have to the (non-normative) world in terms of meaning. Rather, I prefer to say that unlike fundamental statements, non-fundamental statements exhibit mind-to-world fit, owing to their inferential connections to descriptive facts. We can also say that they have truth conditions in a more than minimal sense: they can be said to be true just in case certain features of the world obtain. (Note that all meaningful statements have minimal truth conditions: for example, ‘Happiness is intrinsically good’ is true just in case happiness is intrisically good.) Of course, the more than minimal truth conditions of a given non-fundamental statement differ, in general, from one individual to another, according to their different general attitudes.

Another example of a hybrid theory is that developed by Stephen Barker (Barker 2000). Simplifying slightly, according to his view, if *i* asserts the sentence '*x* is good', then *i* denotes a property *f* by 'good' and expresses as an explicature the content that *x* is *f*, while expressing as an *implicature* the content that *i* is committed to approval of *f*-things. The difference between an explicature and an implicature can be illustrated by means of an example such as:

 Even Granny is drunk

which explicates that Granny is drunk, but implicates that she was one of the most unlikely people to get into that state. The difference is that this unlikelihood – or the belief in it – though it is part of the meaning of the statement broadly construed, is not part of its truth condition. The statement is still going to be true if Granny is drunk even if she is not in fact one of the most unlikely people to get into that state.

Barker in effect sees a normative statement as asserting a descriptive truth condition (in this respect Barker’s theory resembles Edwards’, at least as far as non-fundamental normative statements are concerned). This truth condition is the statement’s explicature. At the same time, the implicative component permits us to acknowledge that normative statements have an attitudinal aspect. Attitudes have been relegated to what a statement conversationally implies, as opposed to the content of the utterance narrowly understood. Is the theory plausible? Barker gets into some difficulty concerning statements about subjects in relation to which our own values differ from that of the speaker. His example involves a case in which a racist called Norm points to an SS officer in a film and says ‘Schmidt was good’. Do we describe Norm’s statement as true or false? Since Norm’s *f* includes being a racist and Schmidt is a racist, it seems we must say it is true. On the other hand, the statement implicates commitment to a certain attitude and part of what that involves, according to Barker, is the ability to treat it as uncontroversial amongst the audience to whom the statement is directed. (The statement *presupposes* this commitment and so it is not open to question.) Of course a commitment to racism is not uncontroversial to us. So we would think it conversationally incorrect to echo Norm’s statement. This, Barker thinks, explains why we might be unhappy to call Norm’s statement true, although literally speaking, it is. However, this strikes me as a rather strained way of saving our intuitions.[[57]](#footnote-57) If possible, it would be better to have a theory which allows us to treat Norm’s statement as straightforwardly false. Understanding a statement ‘*p*’ as false is equivalent to understanding it to be the case that NOT-*p*, and, in Section 3.2, I explained how negations of normative statements, like all other complex statements in which normative statements figure as components, can be understood in an inferentialist way. In fact, one of the examples used involved the relationship between thinking someone a racist and thinking them not good. It was explained how an individual who has a negative attitude to anyone who is a racist could be led to accept:

RACIST2

For any *x*, if *x* is a racist, then *x* is not good.

The important point to note here is that *this individual* should accept RACIST2, even if the context happens to be one in which some racist, such as Norm in Barker’s example, has actually asserted that some particular person is good because they are a racist.

A more recent example of a hybrid theory is Michael Ridge’s ‘ecumenical expressivism’. This theory is rich and subtle and I do not wish to engage with it here in detail. But I would like to say something about Ridge’s theory of truth. He believes that truth is an intrinsically normative concept, supporting this view by pointing out that the primary function of normative discourse is to settle what to do, feel or think and that our use of the truth predicate helps us settle what to think in relation to given propositions (ibid., 206). But granted that, for example, ‘It is true that water is a compound’ could be viewed as trying to settle what to think on the question of whether water is a compound, so could the same statement minus the reference to truth (‘Water is a compound’), which would lead, on Ridge’s reasoning, to a policy of treating all statements as normative, a consequence that would render the concept of normativity pointless. An alternative view is that although ‘It is true that water is a compound’ and ‘Water is a compound’ can both be used to settle the question of whether to think that water is a compound, this is not what either of these statements (and certainly not the latter) *mean*. There are theories of truth according to which this is no part of the meaning of the truth-ascription. One of these is the minimalist theory, and this is the one which accords best with the understanding of the role of normative statements in reasoning defended in Chapter Three.

**CHAPTER SIX**

**Morality: Theory and Rhetoric**

**6.1 The place of theory in ethics**

The radical subjectivism advocated in this book raises a question about the status of moral theories. Given that there is no absolutely objectivity to be had in the moral field (as in any other), and given a conception of the goal of moral theory-construction as seeking to express fundamental moral truths that possess absolute objectivity, it would appear that there is no role for such an activity. However, such a conclusion would be premature. Perhaps it could make sense to seek moral theories even when we know that they will not provide absolute objectivity. The purpose of this section is to explore that possibility.

What are the goals of moral theorising? Why not be satisfied with a mere plurality of moral opinions not presented in any unified form? Here are some possible reasons:

* We may want to justify a moral opinion which has been challenged, and seek theoretical support for this.
* We may want to achieve more agreement about what is morally right or wrong.
* When teaching morality, it is more-or-less essential to use one or more general moral principles, which together at least approximate to a theory.

How are we to go about identifying a sound moral theory? One widely accepted idea is that we should use the method of seeking a *reflective equilibrium*. Originating in the work of John Rawls (Rawls 1999), this term refers to a suggested methodology for doing ethics in which our moral beliefs are weighed against one another in order to achieve maximal coherence within our overall system of beliefs. If certain of our beliefs do not fit in well with a more substantial part of the rest, then they should be jettisoned for the sake of this coherence. When the process has been carried out to the greatest extent possible, the entire set of beliefs that has survived this process, expressed as a set of general principles, constitutes our chosen theory. This process seems likely to fit in with the three goals listed above. It appears to meet the justificatory need, since the justification of a given moral belief would consist in demonstrating its consistency with the general principles, which themselves have been validated by the process of achieving maximal coherence. Also, provided the process does not lead to a multiplicity of different theories when carried out by different people, but rather tends to converge on one theory, or perhaps a small set of very similar theories, the purpose of securing greater agreement will be served. Finally, the general principles, provided they are not too abstract, could presumably form the basis for the teaching of morality.

According to the expressivist/subjectivist view defended in this book, moral beliefs are attitudes based on reasons which count as moral. But at first sight, there is no obvious reason why this should render the three goals mentioned or the method of reflective equilibrium inappropriate. Firstly, the desire for justification does not depend on a wish to achieve absolute objectivity: we may find ourselves wanting justification merely to assuage doubt. Secondly, it may also be important on occasion to achieve greater agreement with others in our attitudes, especially when they concern areas of life where practical cooperation is felt to be needed, and this can be valued without supposing that it would lead to a position which was correct from an absolutely objective standpoint. And thirdly, all societies wish to pass on their fundamental attitudes to their younger members and so need to engage in moral education, quite independently of the ultimate epistemic status of the principles passed on. As for the method of reflective equilibrium itself, given the inferentialist stance adopted in Chapter Three, there seems to be no obvious reason why it should not be applied in cases where what we want is coherence among attitudes rather than descriptive beliefs.

However, despite these considerations, I believe that seeking reflective equilibrium is not a very useful way of doing ethics. An example will help at this point. It concerns a passage from Robert Nozick’s *Anarchy, State and Utopia*. In this passage, Nozick tries to convince us that redistributive taxation is a system of forced labour, with the implication (backed up in other parts of the book) that we should reject it as morally unacceptable:

The fact that others intentionally intervene, in violation of a side
constraint against aggression, to threaten force to limit the alternatives, in this case to paying taxes or (presumably the worse alternative) bare subsistence, makes the taxation system one of forced labor and distinguishes it from other cases of limited choices which are not forcings. (Nozick 1978, 70-71).

It would be possible to construe the process by which Nozick reaches his conclusion as one of seeking reflective equilibrium. One element is what Nozick calls a ‘side-constraint against aggression’. Here aggression is to be understood to include any action which attempts to deprive someone of what they have acquired by just means – in this case, the income derived from selling their labour. Nozick hopes, fairly reasonably, that we accept some general principle to the effect that such actions are wrong. But there is also the fact that many people believe it morally acceptable – even perhaps morally mandated – to impose taxation on individuals with the purpose of giving the money to less economically fortunate members of society. Nozick points out that these two positions are inconsistent. One cannot think it is always wrong to deprive someone of what they have acquired by just means, *and* that redistributive taxation is morally acceptable. Something has to give. If he were to use the language of reflective equilibrium, Nozick would probably argue that the former principle has greater explanatory power than the latter – it accounts for many more specific moral beliefs that we have (such as the wrongness of stealing, extortion, financial fraud and slavery in some of its aspects). Consequently it is the latter that we should jettison. We should give up thinking that redistributive taxation is morally acceptable.

But is this really a sensible way of looking at the matter? Defenders of redistributive taxation are unlikely to find their faith in it shaken by these sorts of considerations. Whether a particular attitude is going to be undermined by considerations of coherence with broader principles will depend on the confidence with which it is held, something which varies from one individual to another. But in that case achieving reflective equilibrium is a rather subjective process which will not further one of the aims identified above, that of achieving greater agreement on moral matters. But going deeper than this, one might ask *why* attitudes which are in a reflective equilibrium should be preferred to ones that are not. And this is a difficult question to answer within the framework of expressivist subjectivism. Reflective equilibrium achieves coherence, which is a kind of simplicity. But why is such simplicity valuable? In the context of descriptive truth, this question is fairly easy to answer. Simpler accounts of the world are more likely to be true than more complicated ones. A simpler theory about who committed a certain crime, requiring few ad hoc assumptions is more likely to be true than a more complicated one involving many such assumptions. In astronomy, the Copernican theory was eventually seen as more likely to be true than the Ptolemaic one even when the latter involved various epicycles introduced to bring it into conformity with observation. Why are simpler theories more likely to be true? The answer seems to be merely that, since this has usually been found to be the case in the past, we can reasonably presume that it will continue to be so in the future. And if it is asked why we should presume that the future will probably resemble the past in this respect, it will likely have to be admitted, as Hume had to admit, that the limit of explanation has been reached. But the point is that these considerations apply to the task of establishing the truth or falsity of descriptive statements. There is no reason to think that they apply to the truth of fundamental normative statements. The use of the word ‘fundamental’ is key here. We saw in Chapter Three that many normative statements have mind-to-world fit: whether they are true or not depends on descriptive features of the world. As a result, it is entirely appropriate to subject them to epistemic procedures in which the value of explanatory simplicity plays a key role. However, what is normally in question when evaluating moral theories is attitudes that are fundamental in that they are not held for reasons and are therefore not subject to mind-to-world fit. In the present case, what is in question is the attitude of favouring redistributive taxation. The issue becomes this: which is worse, the poverty that would be relieved by redistributive taxation or the injustice (as it would be seen by Nozick) involved in it? That is a fundamental attitudinal question, which does not resolve itself into an issue about some feature of the world. There is consequently no reason to apply to it epistemic procedures which are applicable to questions in which mind-to-world fit is the aim of the exercise.

None of this entails that the various attempts by moral philosophers to devise moral theories have been a complete waste of time. They can be of great intellectual interest. Contractarian theory, in particular, presents interesting ideas regarding the empirical question of why we have moral practices at all and why we have certain moral practices and not others. However, moral theories are not, for the most part, useful in respect of the three goals of justification, achieving agreement and moral education. As far as agreement is concerned, it would be optimistic in the extreme to think that this is ever going to be achieved through moral philosophy. Past experience tells a wholly different story! And the theories are usually too abstract and rarefied to be much use for justification[[58]](#footnote-58), still less for moral education. Moreover, what I think is shown by examples like the one from Nozick just explored is that the aim of producing a moral theory which is both completely comprehensive and meets with agreement from all thoughtful individuals will always be unsuccessful. For example, if the theory does not take a stand on the issue of redistributive taxation, then it will not be completely comprehensive, but if it does, its position will be opposed by one or other side in the debate, between whom there is a fundamentally irresolvable disagreement.

**6.2 Objectivity and rhetoric**

It might be thought that since I reject objective values, I cannot accommodate the distinction between objective and subjective ways of arguing for moral conclusions, and that this would be a strong point against my view. But this is not the case. What I deny is that there are *absolutely* objective reasons for things, and absolutely objective ways of arguing. I can nevertheless recognize that that are reasons and ways of arguing that are, to a greater or lesser extent, objective. All arguments, and all reasons offered for anything, if their conclusions are genuinely normative, are subjective in the sense that they are based on the attitudes of the person giving the argument or offering the reason. But an argument or reason can be subjective in a more specific sense to the extent that it is not merely based on the person’s attitudes but favours their selfish interests. An example of this would be a rich person trying to argue against redistributive taxation merely by saying ‘This would hurt me’. The problem with reasons which are subjective in this sense is that they only have force for those who share the same selfish interests.

But undesirable subjectivity can also enter discourse in a more subtle way through the operation of personal bias. For example, a rich person might try to argue against redistributive taxation by claiming that it harms society by disincentivizing wealth-creation, but it could be that the real reason for their attacking this form of taxation is that they see it as detrimental to their selfish interests. In fact, these interests might be the real cause of their belief that it disincentivizes wealth-creation rather than any empirical evidence for this claim. If this is so, then the argument is devalued – those to whom it is addressed have not been given an appropriately *truth-related* reason to believe its premise. Notice, however, that the mere presence of a bias on the part of the person offering the argument does not in itself devalue it. If this person offers empirically sound reasons for thinking that redistributive taxation does disincentivize wealth-creation, then their bias becomes irrelevant. In this lies an important lesson about assessing the significance of personal bias: consider the quality of the argument itself, not the bias (if any) of its author.

Besides the level of objectivity shown in a piece of applied ethics, another salient characteristic is the extent to which it makes use of rhetoric. To illustrate this, we may return to the example of Nozick’s argument against redistributive taxation. Here again is the relevant passage:

The fact that others intentionally intervene, in violation of a side constraint against aggression, to threaten force to limit the alternatives, in this case to paying taxes or (presumably the worse alternative) bare subsistence, makes the taxation system one of forced labour and distinguishes it from other cases of limited choices which are not forcings.

In the discussion above, the argument was viewed in a straightforwardly philosophical way, as an argument to a certain conclusion. But there is another way to see it, namely, as a piece of polemic. In the latter case, the emphasis is on the use of words with emotive colouring, particularly ‘aggression’ and ‘forced labour’.

In traditional ethics, the use of rhetorical language scarcely rates a mention. This is because traditional ethics presents itself as a purely rational enterprise, in which disinterested individuals use cold logic as their modus operandi. But when we study real-life ethical debate, the near-ubiquity of rhetorical expressions is unmistakeable. We all use it to a greater or lesser degree, nor, I would suggest, is there necessarily anything wrong with doing so. On the other hand, most of us have moral attitudes which attach some importance to honesty, and it might be argued that the use of rhetoric is dishonest, and the heavier the rhetoric, the greater the level of dishonesty. When Nozick describes redistributive taxation as ‘forced labour’, he is trying to create in us the emotions that it would be natural to feel in response to what normally goes by that name. Images of the Gulag or similar scenarios rise before our minds, with the attendant condemnatory emotions. Redistributive taxation in modern states, whatever one thinks of its ultimate justifiability, is far more benign than these sorts of practices. It might be countered that Nozick is just giving us his view of the matter, which we are free to accept or reject. But his use of term ‘forced labour’, with its partially descriptive meaning, makes it appear he is merely giving a factual description of redistributive taxation, and since the relevant facts are not in dispute, this can be viewed as somewhat deceptive.[[59]](#footnote-59)

However, in most people’s moral sensibilities, the requirement of honesty is only prima facie and needs to be weighed against other imperatives. The situation may be one in which it is vitally important to persuade someone of the wrongness of a certain act, perhaps because if this is not done, they will perform that act with highly deleterious consequences. In such a case, the importance of preventing the act may outweigh the undesirability of dishonesty. Of course, most situations are far less dire. Can the use of rhetoric be justified in a more run-of-the-mill case? I think we had better say that it can (subject to some qualifications to be discussed below), and for the following reason. A major tool of rhetorical persuasion is the use of thick normative expressions, and, in fact, whenever thick expressions are used, there is a high chance of readers or listeners being influenced by its emotive colouring to take a particular view. The only way of removing this tendency would be to banish thick expressions from our discourse altogether, and this is, of course, an entirely impractical suggestion.

Another important aspect of rhetoric which may influence whether, or to what extent, we should use it is its likelihood of effectiveness. One obvious point is that we should normally expect rhetoric to be less effective on someone who has well-honed logical skills. On the other hand, even the most logically perspicacious of individuals is somewhat susceptible to being swayed by rhetoric if the issue is one on which they do not yet have a firm view.

Another vital question is whether we ourselves should *worry* about being swayed by other people’s rhetoric as opposed to the substance of their arguments. And there are indeed at least two reasons to worry. One is that the use of rhetoric interferes with the operation of descriptive rationality. This is because it tempts us to believe things for emotional reasons, and the beliefs which we may thereby acquire are not necessarily likely to be true, as judged by the standards of descriptive rationality. Clearly this is a drawback, given our preference for true beliefs over false ones. The other reason for concern is very different. It is that we usually value *stability* in our beliefs (normative as well as descriptive). We do not normally like the thought of coming to believe something if we suspect that we are likely to believe the opposite in a few minutes’, or a few hours’, or a few months’, time. This is especially true if the beliefs bear on our own actions, for then such instability can interfere with planning, which, in turn, is likely to frustrate the achievement of our goals. But the dislike of instability may also affect our opinions about matters which do not directly affect our own actions. Clearly, being susceptible to others’ rhetorical persuasiveness can threaten the stability of our beliefs, since our opponents are likely have various rhetorical tricks to call on to which we might be vulnerable at any time.

This concern about some of the consequences of being influenced by rhetoric ourselves, ought, at least if we have any attachment to the Golden Rule (treat others as you would prefer them to treat you), to encourage us in the view that we should usually be restrained in our use of it to persuade others. Though we cannot abolish the practice of rhetorical persuasion, we should be careful about its use. In particular, the most extreme forms of rhetoric should be avoided when our aim is to try to discover, in cooperation with others, what the truth (including the normative truth) on any given matter might be. On the other hand, when our purpose is to try to expose the ineptitude, wickedness or hypocrisy of others, we are entitled to make the most effective use of rhetoric that we can.

**6.3 Seay on theory scepticism and moral dilemmas**

One author who would reject my sceptical approach is Gary Seay. Seay, like most moral philosophers, is an objectivist. He thinks that some moral views are objectively correct, in some absolute sense, while others are objectively incorrect. Broadly speaking, he thinks the correct ones are based on impartial moral principles, which it is the task of moral theory to reveal. He argues that trying to do without some such conception of moral objectivity would make it difficult or impossible to provide convincing resolutions of moral dilemmas (Seay 2002).

I was fairly explicit above about the sense in which I am ‘anti-theory’. Not all the views of the anti-theorists criticised by Seay concerning the nature of morality and moral theory are ones that I would subscribe to. (For example, I do not reject the viability or significance of general moral principles, as some *particularists* such as Jonathan Dancy, do (Dancy 1993). So instead of concentrating on the parts of Seay’s paper in which he discusses these issues in rather general terms, I will concentrate on a specific case in practical ethics in which Seay claims that an opponent of objectively valid moral principles would be left floundering.[[60]](#footnote-60)

The case concerns female genital mutilation (FGM). As many readers will be aware, this has long been common in several parts of Africa. It is severely detrimental to the physical and psychological well-being of the women who are subjected to it, and furthermore, seems integral to a patriarchal system of control over women’s bodies and sexuality. On the other hand, it is embedded in the traditions of the communities concerned; parents think it important, as they believe it is necessary for their daughters’ marriages and even the girls themselves accept it as a way to be ‘part of the group’ (ibid., 292-3).

These points prompt in Seay the thought that a Sudanese father who wishes his daughter to undergo FGM might argue that he can just ‘see’ what is right for her without the benefit of impartial moral principles of the sort which opponents of this practice might appeal to. He compares this to Williams’ example (discussed in the sub-section on universalisation in Section 4.2) of a man appearing to need no further justification for saving his wife from a burning building other than the fact that she is his wife. The problem with this, according to Seay, is that agents may seem to ‘see’ that some quite abhorrent practice - in this case, FGM - is the right thing to do, when an appeal to impartial - and transcultural - moral principles would show this to be mistaken. Thus Seay believes ‘the characteristic doctrines of theory scepticism commit their proponents to positions that in practice will often lead to a kind of impasse in moral reasoning, where persons faced with having to decide what to do in particularly difficult cases are unable to rule out the most implausible conclusions’ (ibid., 295).

Here is my response to Seay’s argument. It is true that we cannot show the Sudanese father’s stance to be mistaken in an absolute sense. Perhaps if we could do this, it would be a good thing. By the same token, it might be a good thing if we had a perpetual motion machine. But since that is quite impossible, we would do best to forget about it. In one sense, this case may not even count as a dilemma for most of us, for we are clear that FGM is wrong. As in Williams’ example, we need no further justification beyond the facts about what FGM involves and its consequences. We do not need a sort of cosmic court which can adjudicate between our stance and that of the Sudanese father. For we do not hold fast to our view because it is right in some absolutely objective sense, but because it is *our* view.[[61]](#footnote-61)

Of course, this is not to deny that there are moral dilemmas. These are cases in which we ourselves do not know what to do. Sometimes reasoning about the facts of the situation--especially facts which may have escaped our notice--can help. But sometimes it doesn’t. There is, sadly, no guarantee of finding a satisfactory resolution, and it is, of course, precisely those cases in which none is to be found, and where the issue is one of great moment, that constitute the most painful dilemmas. However, subjectivism itself may give comfort here. It tells us that in such cases, there is a clear sense in which whatever we do, we are not mistaken. Though this cannot detract from the awfulness of each alternative, it may at least somewhat diminish the mental anguish of choosing one over the others.

**CHAPTER SEVEN**

**Implications of Expressivism**

**7.1 Normative disagreement**

This chapter deals with two miscellaneous aspects of expressivism the issue of normative disagreement and the implications of subjectivism for theism) as well as providing a conclusion to the entire book.

This section is about the vexed subject of whether expressivists can give a convincing account of normative disagreement.

In ordinary factual disagreement, one side asserts a descriptive statement and the other asserts its negation. This is disagreement in belief, where belief is understood in its narrow sense, requiring mind-to-world fit. (Of course, it is disagreement in belief in the broader sense as well.) In contrast, C.L. Stevenson (one of the earliest expressivists) construes normative disagreement as *disagreement in attitude*. He explains that two people disagree in attitude ‘when they have opposed attitudes to something, and when at least one of them is trying to alter the attitude of the other’ (Stevenson 1967, 26). Ridge objects to this theory on the grounds that it does not fit easily alongside what we think about disagreements in belief. For example, if one person believes that the earth is flat and another believes that the earth is round, they disagree even if neither wants to change the other’s belief (Ridge 2015, 171). This might suggest modifying the above analysis so as to omit the clause about one of the people trying to change the other’s attitude. We would then require only that they have opposed attitudes to something. But, as Ridge points out, this theory is vulnerable to counter-examples. For example, we may support different teams in a football match: you want your team to win; I want my team to win. But we do not thereby *disagree* about anything (ibid., 173).

Ridge’s own contribution is to define a notion of ‘disagreement in prescription’ as follows:

Two people (or two stages of the same person at different times), A and B disagree in prescription about D’s φ-ing in C just in case in circumstances of honesty, full candor and non-hypocrisy, A would advise φ-ing in C and B would advise Ψ-ing in C, where φ-ing and Ψ-ing are incompatible, in the sense of being impossible to combine without thereby having inconsistent beliefs (ibid., 190).

It is certainly the case that in the football match example, we do not disagree in prescription according to this definition, which is obviously a point in its favour. But do all normative disagreements count as disagreements in prescription in this sense? A possible counter-example would be a disagreement about whether a certain person is of good character (where the descriptive attributes of the person are agreed on, so that the disagreement is genuinely normative). Ridge suggests that his theory can be extended to cover cases beyond those involving prescriptions about what to do by appealing to the idea of advising what to value, what to count as a reason, and so on (ibid., 192). He would presumably see the apparent counter-example just mentioned as a disagreement about whether to advise morally admiring the person in question. However, this seems a rather strained interpretation. If I say that *a* is a good person, I am not necessarily advising you to admire *a*. I am simply expressing my belief that *a* is in general likely to do morally good things and to avoid doing morally bad ones, and this does not in itself involve advising you to admire *a*.

The question we are faced with is this. When two people express opposing attitudes to something, in the sense that their attitudes could not coherently be held by the same person at the same time, in what circumstances would we say they normatively *disagree*? The initial situation could be a trivial one such as Ridge’s example of two people both wanting a different team to win a football match. Or it could be deeper, such as one person favouring personal happiness as the goal of life and another favouring service to others. The question is: what else needs to be added to the situation so that we would be right to say that there is a normative disagreement between the two people?

It seems to me that there is no one single thing. Instead, there seem to be three distinct things, each of which would make it reasonable to say that a normative disagreement exists.

One is that one of the people involved makes a relevant normative statement ‘*p*’ and the other either makes a statement which explicitly negates ‘*p*’, or a statement which both people believe entails the negation of ‘*p*’ (possibly in conjunction with certain beliefs they both share). For example, in the first of the two above cases, one person says ‘It would be best if Everton won’, and the other says ‘It would be best if Arsenal won’. Again, in the second of the two cases, one person says ‘Personal happiness is the ultimate goal of life’ and the other says ‘Service to others is the ultimate goal of life’. In such cases, it seems hard to deny that there is a genuine normative disagreement between the two people.

We are also likely to say that normative disagreement exists when one or both of the people involved has a second-order attitude which involves judging the other’s first-order attitude to be inferior to their own. In the first case, this would involve one or both people not only wanting their favoured team to win, but also thinking it better to want their team to win than to want the other team to win. In the second case, it would involve not only favouring one of the goals as the ultimate goal of life, but also thinking it better to favour this goal than to favour the other. Notice that in such cases, one person may try to persuade the other to come over to their side, but this is not always the case.

The final thing whose presence would make it reasonable to say that there is a normative disagreement is similar to Ridge’s concept of disagreement in prescription. This would be the situation in which one of the people involved gives certain advice, while the other gives advice which they both believe to be in conflict with it, that is, they both believe it would not be possible to carry out both pieces of advice. This might happen, in the first case, if one person says ‘Support (only) Everton!’, while the other says ‘Support (only) Arsenal!’, and, in the second case, if one person says ‘Seek personal happiness as the ultimate goal of life!’, while the other says ‘Pursue only service to others as the ultimate goal of life!’.

**7.2 God’s values**

If someone is a theist, what are the implications for their belief-system of the claim that there are no absolutely objective values? If this claim is true, then the values God accepts, the values which govern his own actions, and those which he thinks should govern our actions, are not absolutely objective. But this seems like an uncomfortable view for the theist to adopt. It appears to imply that God’s values are arbitrary. He could have chosen a completely different set of values, perhaps ones based on what we think of as evil, and there would have been nothing wrong in his doing so in an absolutely objective sense.

It might be objected that God could not in fact have chosen any values he liked, since our very concept of God is the idea of a person who has such traits as being all-loving and all-virtuous. Even the idea of God’s *choosing* one set of values over another is incorrect, according to this view. A being does not get to choose what is part of their essence. That it is impossible for God to have had different values is not incompatible with God’s omnipotence, it might be claimed, because to suppose that God had some different set of values from those he actually has would be a contradiction, and omnipotence does not include the ability to bring about a contradiction. However, the argument is flawed for a different reason. For another property that God has is that of being creator of the Universe, and there is no contradiction in supposing that the person who is the creator of the universe might have had different values from those he actually has. Hence the argument that it would be impossible for God to have had different values because these values are entailed by the concept of God fails to establish its conclusion.

The fact that God could have chosen different values might not be a great problem for the theist if the values he actually has are ones that we can respect. In one area, this may be true. If we consider the Christian God and invoke the New Testament, we do find there some values, such as those expressed in the Sermon on the Mount, to which it is easy to feel sympathetic. However, when we come to the values that appear to govern God’s own choices, as opposed to those which he has supposedly ordained us to follow, it is not always easy to feel the same sympathy. I have in mind the fact that the world contains a great deal of suffering. Some of this suffering is caused by human actions. An omnipotent God could prevent this, but chooses not to. Perhaps the most common response made by Christian apologists to this problem is that God wants humans to have the ability to make free choices to do good, but that this is not possible unless there is a chance of our choosing instead to do evil, with its frequent attendant suffering. But while most of us would not deny that there is some value in freely chosen good actions, it is hard to accept that it is enough to outweigh the terrible suffering caused by human wickedness. And there is, in addition to this, the suffering which is not caused by human actions, such as that due to natural disasters. Perhaps, as so-called sceptical theists maintain, these further some divine purpose which it is hard for us, with our limited minds, to fathom. But this response is less effective than is sometimes thought. The argument seems to be that if we had God’s view of the matter, we would appreciate the purposes for which the existence of evil is necessary, together with the fact that these purposes are indeed sufficient to justify its being brought into existence, or at least its being permitted to exist. However, this argument does not work if one is a normative subjectivist. Normative *objectivism* might allow one to treat the claim that permitting the evil is justified as a descriptive claim, and thus to construe the proposed argument as saying that there could be evidence for this claim which our minds are not able to understand, so that if were able to understand it, we would appreciate its force. This argument seems reasonable, but without normative objectivism such a construal is not possible.[[62]](#footnote-62) Instead, it would amount to no more than the claim that if a certain state of affairs obtained, we would make different normative judgements. This is true, but it does not in any way undermine the normative judgements that we actually do make.

Does this show that the existence of evil for which there could be no justification refutes theism? This cannot be given a definitive answer, as whether there exists a being which is both all-powerful and all-good depends on whether there is sufficient justification for an all-powerful being to permit the evil that exists, and it may be that not everyone agrees with me that the evil exists is so terrible that there could no be sufficient justification for it (some may, for example, be convinced by the free will defence). For those who do agree with me about this, there cannot be a God, since there cannot be a being who is both all-powerful and all-good. Of course, this does not entail that there is no being who is all-powerful who has values very different from my own. Obviously, I hope that this is not the case, but whether it is or not can only be determined (if at all) by other considerations which go beyond the scope of this book.

While we are on the subject of God, though, it is worth mentioning a possible misunderstanding of the claim that there are no absolutely objective values. One might think that the values of God would be, by definition, absolutely objective. After all, there is surely no greater authority than God. But this would be tantamount to accepting that whatever values God has are absolutely objectively right, and hence that if God were to desire unending misery for all his creatures, that would be an absolutely objectively desirable thing. But this seems absurd. For this reason, it seems that Christians are forced to accept the monotheistic equivalent of the view explored in Plato’s *Euthyphro* that the gods love the pious because it is pious, rather than the reverse (Plato). Since we cannot therefore define the absolutely objectively valuable as that which God values, the concept of the absolutely objectively valuable remains as mysterious as ever.

**7.3 Conclusion**

According to expressivism, normative statements expreess attitudes. This includes a statement such as:

MARS1

If your aim is to believe only what is true, you should believe there is no life on Mars.

It might be thought that this is not genuinely normative, but more like what Kant called a hypothetical imperative (equivalent perhaps to believing that there is no life on Mars is necessary for believing only what is true) and hence a descriptive statement. But I think that is mistaken. Having the aim of believing what is true does not in itself make a person believe that there is no life on Mars. MARS1 expresses a positive attitude to one particular possible outcome of addressing the question of whether there is life on Mars. Whatever answer is given, the opposite answer would have been possible and would not have been mistaken in an absolutely objective sense.

But if this is right, it can seem that the fact/value distinction is not so important after all. For anyone who accepts MARS should also accept

MARS2

There is no life on Mars

which is clearly a descriptive statement. So while the disttinction between factual (i.e., descriptive) statements remains technically valid, it does not, as is normally thought, create a cleavage between matters for which absolute objectivity is possible and matters for which it is not.

There *is* a distinction to be made, though it is a purely contingent one. In factual matters, even when people disagree, it is oten possible in principle to find a way of resolving the disagreement (e.g., go to Mars and look). But in fundamental normative matters this is rarely possible. This, then, is why the fact/value distinction is important.

In one way, the main claims of this book are very radical. Many, perhaps most, moral philosophers still cling to the idea that there are fundamental objective moral values underlying our moral judgements. They will reject the Redundancy Thesis, as it seems outrageous to suggest that normative statements contribute to our discourse nothing more than surface features, such as succinctness, elegance and rhetorical force.

But there is one very important thing to notice about this radicalism, which is that these views are only radical in the context of moral philosophy (and, to some extent, the theory of knowledge). As far as everyday life is concerned, almost everything can proceed as normal. And this includes the part of everyday life that is concerned with ethical matters and with intellectual matters other than the meta-ethical and meta-epistemic matters dealt with in this book. People will still disagree with others about a range of issues, and will often demand that these others justify their views. The Redundancy Thesis is not radical with respect to our normal ethical discourse, since, although it tells us that we *could* remove all normative claims from our discourse, it does not say that we *should*. And despite the fact that there are no absolutely objective values, people will continue to forcefully assert their positions on many issues, and quite rightly in many cases, as normative subjectivism in no way requires us to weaken our level of commitment to our strongly held normative views. In opposing ethical relativism, Simon Blackburn, a prominent expressivist, has had this to say:

[The relativist] will say [concerning an ethical disagreement]: “well, it is *merely* your attitude against his”. Part of this is right: it is indeed my attitude against his. That is what ethical conflict is. The part that is wrong is the “merely”. What is “mere” about a conflict of attitude? The world’s worst conflicts are those of policy, choice, and practice. They are the most important conflicts there are. (Blackburn 2000, 40)

In this book, I myself have tended to say that normative statements are *merely* expressions of attitude. This is because many philosophers appear to think that they are something more, and that seems to me incorrect. But I agree with Blackburn that in the context of everyday life, there is nothing ‘mere’ about attitudes.

**GLOSSARY**

*Not all the technical terms used in this book are defined here. I have confined myself to terms which are either (a) crucial to the understanding of expressivism or (b) non-standard.*

**Assertability condition of a statement**

The condition under which it is semantically permissible for a speaker to assert the statement (Schroeder 2008, 31). For example, the assertability condition of ‘Water is a compound’ is that the speaker should believe that water is a compound.

**Normative statement**

A statement whose assertability condition is that the person making it has a certain attitude. In any context in which the statement functions so as to make an assertion (e.g., when stand-alone, or in a conjunction) the statement is also said to *express* that attitude.

**Expressivism**

The philosophical position according to which there is nothing to the concept of a normative statement which is not brought out by the above definition.

**Descriptive statement**

A statement which is not normative.

**Normative term**

A term (word or phrase) whose inclusion in a statement makes the statement normative.

**Descriptive term**

A term (word or phrase) which is not normative.

**Thick normative term**

A normative term which has a descriptive aspect to its meaning, i.e., any statement in which it occurs entails a descriptive statement in virtue of the meaning of that term (for example, ‘Hitler was cruel’ entails ‘Hitler caused others suffering’ in virtue of the meaning of the word ‘cruel’).

**Thin normative term**

A normative term which is not thick.

**Pure normative statement**

A normative statement which does not entail any descriptive statements. The only normative predicates in such a statement will be thin (e.g. ‘right’ or ‘good’).

**Speaker subjectivism**

The thesis according to which a normative statement describes the speaker’s own attitudes.

**Redundancy thesis**

The thesis according to which if we did not make any normative statements, but instead used only descriptive terms in our statements (with reference, where necessary, to our own subjective attitudes), the only losses, in general, would be, at most, stylistic variety, rhetorical force, elegance and/or succinctness.

**BIBLIOGRAPHY**

|  |  |
| --- | --- |
| Ahmed, Arif 2017. | “Exploiting Cyclic Preference”, *Mind* 126(504): 975-1022 |
| Ayer, Alfred Jules 1946. | *Language, Truth and Logic*, Victor Gollancz (London) |
| Bambrough, Renford 2021. | *Moral Scepticism and Moral Knowledge*, Routledge (London & New York) |
| Barker, Stephen J. 2000. | “Is value content a component of conventional implicature?”, *Analysis* 60(3): 268-279 |
| Blackburn, Simon 1984. | *Spreading the Word*, Oxford University Press (Oxford) |
| ------------------------ 1988. | “Attitudes and Content”, *Ethics*, 98(3): 501-517, reprinted in Shafer-Landau & Cuneo 2007, 474-484. |
| ------------------------ 1998. | *Ruling Passions: A Theory of Practical Reasoning*, Oxford University Press (Oxford) |
| ------------------------ 2000. | “Relativism” in LaFollette (2000): 38-52 |
| ------------------------ 2010. | *Practical Tortoise Raising*, Oxford University Press (Oxford, New York) |
| Boisvert, Daniel R. 2008. | “Expressivism-assertivism”, *Pacific Philosophical Quarterly* 89(2): 169-203 |
| Boyd, Richard N. 2007. | “How to be a moral realist” in Shafer-Landau and Cuneo 2007, 474-484 |
| Brandom, Robert B. 2000. | *Articulating Reasons: An Introduction to Inferentialism*, Harvard University Press (Cambridge Massachusetts; London) |
| Brandt, Richard B. 1979. | *A Theory of the Good and the Right*, Oxford University Press (Oxford) |
| Brook, Richard & Schwimmer, Seymour 1981. | “On Adding the Good”, *Social Theory and Practice*, 7 (3): 325-335 |
| Edgington, Dorothy 1995. | “On Conditionals”, *Mind*, 104(414): 235-329. |
| Edwards, Paul 1955. | *The Logic of Moral Discourse*, Free Press (Glencoe, Illinois) |
| ----------------------- 1968.  | “Bertrand Russell’s Doubts about Induction” in Flew 1968, 55-79 |
| Flew, A.G.N. (ed.) 1968. | *Logic and Language* (first Series), Blackwell (Oxford) |
| Foot, Philippa (ed.) 1967. | *Theories of Ethics,* Oxford University Press (Oxford) |
| ------------------------- 2002. | *Virtues and Vices and Other Essays in Moral Philosophy*, second edition (first edition 1978), Oxford University Press (Oxford) |
| Gauthier, David 1986. | M*orals by Agreement*, Oxford University Press (Oxford) |
| Geach, P.T. 1965. | “Assertion”, *Philosophical Review* 74(4): 449-465 |
| -------------- 1967. | “Good and Evil” in Foot 1967: 64-73 |
| Gibbard, Allan 1990. | *Wise Choices, Apt Feelings: A Theory of Normative Judgement*, Oxford University Press (Oxford) |
| Hare, R.M. 1952. | *The Language of Morals*, Oxford University Press (Oxford) |
| -------------- 1967.  | “The Promising Game” in Foot 1967: 115-127 |
| -------------- 1981. | *Moral Thinking: Its Levels, Method and Point*, Oxford University Press (Oxford) |
| Harsanyi, John 1982. | “Morality and the Theory of Rational Behaviour” in Sen & Williams 1982: 39-62 |
| Horwich, Paul 1994. | “The Essence of Expressivism”, *Analysis* 54(1): 19-20. |
| ------------------ 1998. | *Truth*, Oxford University Press (Oxford). |
| Hume, David 1911. | *A Treatise of Human Nature* (2 volumes). J.M. Dent and Sons (London) |
| Hyde, Dominic 1997. | “From Heaps and Gaps to Heaps of Gluts”, *Mind* 106(424): 641-660 |
| Kant, Immanuel 1898. | “On a Supposed Right to Tell Lies from Benevolent Motives” in T.K. Abbott (trans.), *Kant's Critique of Practical Reason and Other Works on the Theory of Ethics*. Longmans, Green and Co. (London) |
| –-------------------- 1956. | *The Moral Law or Kant's Groundwork of the Metaphysic of Morals* (trans, H.J. Paton), Hutchinson & Co. (London) |
| Kölbel, Max 2007. | “How to Spell Out Genuine Relativism and How to Defend Indexical Relativism”, *International Journal of Philosophical Studies* 15(2): 281–288 |
| Kripke, Saul 2011. | “Philosophical Troubles” (Collected papers vol. 1), Oxford University Press (Oxford) |
| Kuhn, Steven 2019. | “Prisoner’s Dilemma”. *Stanford Encyclopedia of Philosophy*. (Winter 2019 Edition), Edward N. Zalta (ed.), [<https://plato.stanford.edu/archives/win2019/entries/prisoner-dilemma/>](https://plato.stanford.edu/archives/win2019/entries/prisoner-dilemma/) (retrieved 18 March, 2021) |
| LaFollette, Hugh (ed.) 2000. | *The Blackwell Guide to Ethical Theory*, Blackwell (Malden MA, Oxford) |
| Lajevardi, K 2016. | “Kripke and the dogmatism paradox”, unpublished paper |
| Leslie, John 1991. | “Ensuring Two Bird Deaths With One Throw”, *Mind* 100(1): 73-86 |
| Mackie, J. L. 1985. | *Persons and Values: selected papers vol. 2*, Oxford University Press (Oxford) |
| --------------- 1990. | *Ethics: Inventing Right and Wrong*, Penguin Books (London) |
| McDermott M. 1982.  | “Utility and Distribution”, *Mind* 91 (364): 572-8 |
| McDowell, John 1998. | *Mind, Value and Reality*, Harvard University Press (Cambridge, MA) |
| Moore, G.E. 1903. | *Principia Ethica*, <https://www.gutenberg.org/files/53430/53430-h/53430-h.htm> (retrieved 8 July, 2024). |
| Nagel, Thomas 1986. | *The View From Nowhere*, Oxford University Press (Oxford) |
| Nozick, Robert 1969.  | “Newcomb’s Problem and Two Principles of Choice”, in R. Rescher *et al.* (eds.) *Essays in Honor of Carl. F. Hempel*, Dordrecht (Reidel): 114–146 |
| –------------------ 2001. | *Anarchy, State and Utopia*, Blackwell (Oxford) |
| Parfit, Derek 1984. | *Reasons and Persons*, Oxford University Press (Oxford; New York) |
| ----------------- 2011-2017. | *On What Matters* (3 volumes), Oxford University Press (Oxford) |
| Plato. | *Euthyphro,* [*http://classics.mit.edu/Plato/euthyfro.html*](http://classics.mit.edu/Plato/euthyfro.html)(retrieved 19/11/22) |
| Putnam, Hilary 1981. | *Reason, Truth and History*, Cambridge University Press (Cambridge) |
| Rai, Tage S. & Holyoak, Keith J. 2013. | “Exposure to moral relativism compromises moral behaviour”, [Journal of Experimental Social Psychology](https://www.researchgate.net/journal/Journal-of-Experimental-Social-Psychology-1096-0465) 49(6):995–1001 |
| Railton, Peter 2007. | “Moral realism” in Shafer-Landau (2007): 186-205 |
| Rawls, John 1999. | *A Theory of Justice* (revised edition), Harvard University Press (Cambridge, MA) |
| Ridge, Michael 2006. | “Sincerity and Expressivism”, Philosophical Studies 131(2): 487 – 510 |
| ------------------- 2015. | *Impassioned Belief*, Oxford University Press (Oxford). |
| Robert, David 2018. | “A restatement of expected comparative utility theory: A new theory of rational choice under risk”, [*Philosophical Forum*](https://philpapers.org/asearch.pl?pub=782) 49 (1): 19-37. |
| Rosen, Gideon 2004. | “Skepticism about Moral Responsibility”, *Philosophical Perspectives*, vol. 18, Ethics: 295-313 |
| Ross, David 2002. | *The Right and the Good* (ed. Philip Stratton-Lake), Oxford University Press (Oxford) |
| Rumfitt, Ian 2000. | “ 'Yes' and 'No' ”, *Mind* 109(436), 781-823. |
| Sayre-McCord, Geoff 2021. | "Moral Realism", *The Stanford Encyclopedia of Philosophy*(Summer 2021 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2021/entries/moral-realism/> |
| Schroeder, Mark 2008. | *Being For*, Oxford University Press (Oxford). Kindle Edition |
| Searle, John R. 1967. | “How to derive ‘ought’ from ‘is’” in Foot, 1967: 101-114 |
| Seay, Gary 2002. | “Theory Skepticism and Moral Dilemmas”, *Kennedy Institute of Ethics Journal* 12(3): 279-298 |
| Sen, Amartya & Williams, Bernard (ed.) 1982. | *Utilitarianism and Beyond*, Cambridge University Press (Cambridge); Editions de la Maison des Sciences de l’Homme (Paris) |
| Shafer-Landau, Russ 2007. | “Ethics as Philosophy” in Shafer-Landau and Cuneo 2007 |
| Shafer-Landau, Russ & Cuneo, Terence (ed.) 2007. | *Foundations of Ethics: An Anthology*, Blackwell(Malden MA, Oxford, Victoria) |
| Simmons, Howard 2010. | *Moral Desert: A Critique*, University Press of America. (Lanham, Maryland; Plymouth, U.K.) |
| Singer, Peter 1993. | *Practical Ethics*, Cambridge University Press (Cambridge) |
| Smith, Michael 1994. | *The Moral Problem*, Blackwell (Oxford) |
| -------------------- 2000. | “Moral realism” in LaFollette 2000: 15-37 |
| Sorell, Tom 2000. | *Moral Theory and Anomaly*, Blackwell (Oxford; Malden, Massachusetts) |
| Steele, Katie & Stefánsson, H. Ori 2020. | "Decision Theory", *The Stanford Encyclopedia of Philosophy*(Winter 2020 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2020/entries/decision-theory/> |
| Stevenson C.L. 1944. | *Ethics and Language*, Yale University Press (New Haven) |
| ------------------- 1967. | “Moore’s Arguments against Certain Forms of Ethical Naturalism” in Foot (1967): 16-32 |
| Strawson, Galen 1994.  | “The Impossibility of Moral Responsibility”, *Philosophical Studies*, 75(1–2): 5–24 |
| Thomson, Judith Jarvis 2008. | *Normativity*, Open Court Press, (Chicago) |
| Unger, Peter 1975 | *Ignorance: A Case for Scepticism*, Oxford University Press (Oxford) |
| Van Roojen, Mark 2005. | “Expressivism, Supervenience and Logic”, *Ratio* 18(2): 190-205 |
| Williams, Bernard 1981. | *Moral Luck*, Cambridge University Press (Cambridge) |
| ------------------------ 2011. | *Ethics and the Limits of Philosophy*, Routledge (London & New York) |
| Wittgenstein, Ludwig 1961. | *Tractatus Logico-Philosophicus* (translated by D.F. Pears & B.F. McGuinness), Routledge (London) |
| ---------------------------- 1963. | *Philosophical Investigations* (translated by by G.E.M. Anscombe), Blackwell (Oxford) |
| Yalcin, Seth 2012. | “Baysian Expressivism”, *Proceedings of the Aristotelian Society*, New Series vol. 112, 123-160 |

1. I use the word ‘statement’ to mean a sentence which we would pretheoretically describe as true or false, e.g., ‘Paris is the capital of France’ or ‘Rain!’, when the latter is intended to state that rain is falling. Notice that a statement may be a statement type or a statement token produced on a given occasion. I use the word ‘proposition’ to refer to what is believed, imagined, supposed etc. and which can be represented by a statement. [↑](#footnote-ref-1)
2. Throughout this book, when I make a claim about the meaning of a certain (type of) statement or expression, I intend to be taken as saying that this claim applies at least to ‘standard’ or ‘normal’ uses of it. [↑](#footnote-ref-2)
3. This applies to simple, ‘stand-alone’ statements only. As we will see, complex statements in which simple normative statements are embedded, are a different matter. [↑](#footnote-ref-3)
4. For this reason, we may observe that Hilary Putnam’s attempt to undermine the fact/value dichotomy by noting that ‘one cannot apply a word at all if one places oneself entirely outside of the tradition to which it belongs’ (Putnam 1981, 203) is really a straw man argument. For the defender of the dichotomy need not suppose that speakers are ever able to ‘place themselves entirely outside of the tradition’, only that they are capable of acting within the tradition without necessarily endorsing it. [↑](#footnote-ref-4)
5. It might be argued that the mere act of making any statement, being voluntary, reveals at least the speaker’s positive attitude to the making of that statement at that moment. But here again, we can simply point out that even if this is accepted, the attitude involved, not being related to the subject-matter of the statement, falls outside of what we here intend by the term ‘normative’. [↑](#footnote-ref-5)
6. The italicised term is Mark Schroeder’s (Schroeder 2008, 4). [↑](#footnote-ref-6)
7. A close cousin of speaker subjectivism is the theory Mark Ridge calls *naïve subjectivism*, according to which normative statements describe the speaker’s feelings or emotions. This is subject to problems similar to those discussed here. See Ridge 2015, 60-64. [↑](#footnote-ref-7)
8. A classic example is McDowell 1998,154-169. [↑](#footnote-ref-8)
9. Examination of the history of this article’s edits shows that in an earlier version, the word ‘shackles’ was used rather than ‘restrictions’. This seems like a *clear* case of normative negativity, which, as an editorial note (‘more neutral wording’) reveals, was the reason for making the substitution. [↑](#footnote-ref-9)
10. C.L. Stevenson shows, through his concept of a ‘persuasive definition’, how this fact can be harnessed by speakers to proselytise for their own attitudes (Stevenson 1944, Chapter 9). [↑](#footnote-ref-10)
11. Words like ‘must’ and ‘essential’ appear to be exceptions to this. For example, it is incoherent to say ‘You must do it, but I have no attitude regarding the prospect of your doing it’. [↑](#footnote-ref-11)
12. The use of the word ‘categorically’ here and elsewhere in this paraphrase is intended to convey the idea that the authors would not want anything to override the stated desire. This seems a fair description of the mental state expressed by the use of the word ‘must’. [↑](#footnote-ref-12)
13. Some writers maintain that assertions are subject to a more stringent requirement, e.g., that the speaker should know the proposition to be true. (See, e.g., Unger 1975.) I do not necessarily disagree with such suggestions. Believing the proposition is meeting a low bar, but it is meeting a bar nonetheless. [↑](#footnote-ref-13)
14. Or a certain set of attitudes in more complex cases. [↑](#footnote-ref-14)
15. Such misspeaking could be the result of believing one’s attitude to be different from what it is, or it could have a more banal explanation, namely that one does not understand the correct use of this sort of statement. Note that according to my terminology, in such cases the statement expresses an attitude which the speaker does not have, but if this is thought to be too strange, one could alternatively say that the attitude is expressed only when the speaker has the attitude and understands that the statement should only be made if that is the case. [↑](#footnote-ref-15)
16. It has been pointed out that this way of drawing the distinction is faulty, since it implies that the idea of something’s being big full stop makes sense, which it does not (bigness is always relative to some standard). Judith Jarvis Thomson later gave an improved definition by requiring that for an adjective ‘*a*’ to be attributive, being an *a n* and being an *m* must not entail being an *a* *m* for any arbitrary *m* (Thomson 2008, 3-5). [↑](#footnote-ref-16)
17. There is also perhaps the implication that the action is possible under the circumstances, though this seems to be a presupposition of the statement, rather than part of what it asserts. [↑](#footnote-ref-17)
18. Not everyone accepts that deontic statements have motivational load. Reasons for scepticism will be discussed in Section 5.2. [↑](#footnote-ref-18)
19. Thus the assertability condition of ‘Simon ought not to have done that’ is the action’s belonging to a type of which the speaker disapproves. (It is not that the speaker believes this to be the case, as that would make facts about the speaker’s attitudes part of the content of what is asserted, creating a theory which would be akin to the speaker subjectivist view criticised in Section 1.1.) [↑](#footnote-ref-19)
20. Throughout this paragraph, the term ‘action’ means an action type or an individual action. [↑](#footnote-ref-20)
21. Ridge usefully emphasises the importance of emotions in stabilising our normative commitments (Ridge 2015, 132). [↑](#footnote-ref-21)
22. In saying this, I leave aside the possibility that any non-humans think in ‘mentalese’. [↑](#footnote-ref-22)
23. Actually, this would be slightly too restrictive. Sometimes embedded normative statements do express attitudes, as in: ‘He acted courageously, and he should be proud of that fact’. It depends on the meaning of the relevant connective (in this case ‘and’) whether the embedded statement is asserted or not. [↑](#footnote-ref-23)
24. If the argument is non-demonstrative, we accept that its conclusion is *probably* true if its premises are, so truth is involved here too. [↑](#footnote-ref-24)
25. Surprisingly, this has to be regarded as a defeasible principle, not an absolute one, as is shown by this counter-example from Dorothy Edgington: ‘If Reagan was in the pay of the KGB, we’ll never find out’ (Edgington 1995, 270). On pain of absurdity, believing this conditional would not dispose me to believe that we’ll never find out that Reagan was in the pay of the KGB if I come to believe that he was. But this case is very much an anomaly. It has the peculiar property that my belief in the antecedent would undermine any belief I might have in the conditional, since to believe both would amount to a kind of Moorean inconsistency, similar to that between ‘There is life on Mars’ and ‘I do not believe there is life on Mars’. [↑](#footnote-ref-25)
26. Compare the idea in Blackburn 1988 that when we accept a disjunction, whether or not its disjuncts are descriptive or normative, we may be said to be ‘tied to a tree’ (that is to say, we are tied to accepting one disjunct should the other prove to be untenable). But his development includes a semantic theory for normative statements. (An alternative semantic theory is presented in Gibbard 1990.) If I am right, then no such theory is necessary. Blackburn’s view has evolved, though. In later work, ‘he derives propositionality from function more directly, and with less commitment to any kind of contrast with description or representation’ (Blackburn 2010, 46). [↑](#footnote-ref-26)
27. For a general defence of the inferentialist approach, see Brandom 2000. [↑](#footnote-ref-27)
28. For a general defence of this view, see Horwich 1998. [↑](#footnote-ref-28)
29. There are exceptions, of course. If I point outside and say ‘Look, rain!’ my utterance could legitimately be described as true or false. But in such cases, the sentence is always equivalent to one expressed in the grammatical form which is more standard for truth-aptness (‘It’s raining outside’, say). [↑](#footnote-ref-29)
30. The term ‘indicative’ can be used loosely to refer to the correct grammatical form for truth-aptness, though it isn’t quite accurate, as subjunctive statements are also truth-apt. (The intended contrast is with imperatives, which are definitely not truth-apt.) [↑](#footnote-ref-30)
31. I am not trying to suggest that this scenario somehow *refutes* the point that Dreier is making. He is merely pointing out that the mere fact that it is possible to be a minimalist about truth-aptness does not in itself explain how expressions of attitude realised as declarative sentences would actually function in all the ways that we would need them to. My account here is intended to provide that explanation. [↑](#footnote-ref-31)
32. There are also conditionals involving normative words which are analytic, such as ‘‘If all lies are wrong, then John’s lie was wrong’ and ‘If this is good and that is bad, then this is better than that’. It would be hard to give a full explanation of these without invoking a general theory of analytic truth, which would go far beyond the scope of this book. But at any rate, there is no obvious reason for thinking that our account of analytic truth would be affected by the presence of normative words. [↑](#footnote-ref-32)
33. It might be objected that this account overlooks the moral fervour with which such attitudes are typically held. There is a difference, it will be suggested, between not *wanting* a statue to be there and thinking it morally wrong for the statue to be present. I will argue in the next chapter that moral attitudes are those formed on the basis of certain sorts of considerations rather than others. That aside, I think the difference is simply a matter of the strength of the desire or attitude involved. [↑](#footnote-ref-33)
34. Throughout this discussion, except where otherwise stated, I assume the normative statement or proposition involved is pure, i.e., that it is not a mixed normative statement or proposition. [↑](#footnote-ref-34)
35. Besides analytic beliefs about what drinking alcohol without restraint consits in. [↑](#footnote-ref-35)
36. Another expressivist who explicitly rejects moral relativism is Simon Blackburn (Blackburn 2000). [↑](#footnote-ref-36)
37. This distinction is also made by G.J. Warnock (Warnock 1967, 52-5), though his terminology is different, as is his precise list of structural features (though there is certainly overlap with mine). In fact, his list is strictly a list of features of morality, as understood by ordinary people, and does not contain elements (such as the impossibility of moral conflicts) suggested by moral *philosophers* (see also note 2). [↑](#footnote-ref-37)
38. The features are those identified in Williams’ *Ethics and the Limits of Philosophy* (Williams 2013, 193-218), (though he would not necessarily have agreed with my various comments on them) plus one (the last one about centrality) based on an idea of Blackburn’s (Blackburn 1998, 67). One feature discussed by Williams which I have not included here is the idea that ‘ought implies can’, as this is not specifically about morality. More generally, the reader will notice some key differences between my critique and that of Williams. He talks of the ‘peculiar institution of morality’, though much of what he has to say concerns not morality as it is understood by ‘ordinary people’, but moral *theory*, a more rarefied (and second order) thing. [↑](#footnote-ref-38)
39. According to Hare, moral judgements form a special class of *prescriptive* judgements. His concept of prescriptivity is similar to the implicit advice-giving property attributed to deontic statements in Section 2.3. Roughly speaking, moral judgements are for Hare those prescriptive judgements which are both universalisable and overriding (I will discuss the overridingness property in the next subsection). [↑](#footnote-ref-39)
40. I also assume that the martyrs know fully and accurately the fate that awaits them. [↑](#footnote-ref-40)
41. We see something analogous to this also in cases where imperialists suppose that normal moral protections do not apply to the inhabitants of certain other countries. Though grotesque, this is surely an intelligible position. [↑](#footnote-ref-41)
42. Clashes between moral intuitions and moral theory raise the subject of reflective equilibrium, which is discussed in Section 6.1. [↑](#footnote-ref-42)
43. Strictly speaking, we could formulate such conditions if we allowed them to be very complex and somewhat arbitrary, but this would not satisfy the aim usually motivating such attempts, which is to find a genuine ‘common property’ (Wittgenstein 1963, sec. 67 (last paragraph), 32). [↑](#footnote-ref-43)
44. This set of conditions can be used to define prima facie or decisive moral wrongness according to whether the negative attitude involved is prima facie or decisive. [↑](#footnote-ref-44)
45. I am assuming here that ‘good’ is being understood in a general normative sense. If it means ‘morally good’, then the analysis is not subject to this objection, but would appear to be vulnerable to a different objection, namely, that it seems to involve an identification of morality with rationality, which is dubious. [↑](#footnote-ref-45)
46. A version of this argument appears in Smith 2000, 27. He thinks it ‘constitutes a decisive refutation of expressivism’. I think this is unwarranted, for the reason explained here. [↑](#footnote-ref-46)
47. For simplicity, I here ignore the point raised in Chapter Two about ‘good’ being an attributive adjective. The proposed schema could easily be modified to accommodate this. [↑](#footnote-ref-47)
48. An example of a normative statement which, by itself, does not have this property might be ‘McArthur was wrong to advance into North Korea in 1950’. Unless the speaker adds *why* she thinks this, in other words what general rule she thinks McArthur was violating, there is no motivational load incurred for her. (Even then, the motivational load may be of an extremely hypothetical character, as she might be very unlikely ever to find herself in a situation where the rule applies to her.) [↑](#footnote-ref-48)
49. Some other studies are discussed in Ridge 2015, 51-4. [↑](#footnote-ref-49)
50. Strandberg and Björklund 2011, 8. [↑](#footnote-ref-50)
51. To be fair, Strandberg and Björklund themselves are cautious. They recognise some methodological limitations in their study, for example, the possibility that the word 'could' may not have been understood by the subjects in the intended sense. Their conclusion is that 'the results of the study provide reasons to doubt the evidential basis for claiming that there is a non-trivial conceptually necessary connection between moral judgments and motivation' (Strandberg and Björklund 2011, 16).

 [↑](#footnote-ref-51)
52. This is of course assuming the naturalistic (and hence descriptive) reading of the word ‘rational’. We saw earlier that a normative reading would be possible, but would be self-defeating for a naturalist. [↑](#footnote-ref-52)
53. This leaves unaddressed Railton’s emphasis on the impartial point of view. But we discussed that notion in Section 4.2 under the guise of universality. It will also re-emerge in a different form in Section 6.7 when we look at the question of objectivity. [↑](#footnote-ref-53)
54. Boyd uses the word ‘defined’ rather than ‘identified’, but he makes it clear elsewhere that he is not talking about the sort of definition which describes the conventional uses of terms, but ‘natural’ definitions (ibid., 170). [↑](#footnote-ref-54)
55. He consequences of accepting this are briefly explored in the last sectionof the book. [↑](#footnote-ref-55)
56. That no genuinely normative statement follows from any set of descriptive statements is convincingly defended in Hare 1967 in response to Searle 1967. [↑](#footnote-ref-56)
57. To be fair, Barker does suggest an alternative interpretative strategy which would enable Norm’s sentence to come out false. He says that as Norm is signalling that his attitude is shared by or uncontroversial for the intended audience, we can assign one of our own attitudes to Norm's utterance. Then we will take the relevant property to be linked to our conception of goodness, and so Norm’s statement will come out as false, since Schmidt lacks that property (ibid., 278). The trouble is that the presupposition of certain shared attitudes is not present in every case. Sometimes people say that things are good or bad, or actions right or wrong, knowing full well that their audience will not agree with the attitudes which they thereby express. [↑](#footnote-ref-57)
58. The broadly contractarian thought ‘If you do not follow morality, you’ll be the loser’ might be thought to justify morality, but even to the extent that it is true, it is too obvious to count as a moral *theory*. [↑](#footnote-ref-58)
59. To be fair to Nozick, it should be mentioned that in a footnote, he wonders whether it is really correct to say that redistributive taxation is (a kind of) forced labour or whether it is merely ‘on a par with’ forced labour. But my quotation occurs after the reference to this footnote, and so I think it is fair to take it as representing Nozick’s considered view. [↑](#footnote-ref-59)
60. Seay also discusses another case, concerning cochlear implants, but I will leave this one aside, as he specifically deals with it as a problem for particularists, understood as rejecting the viability or usefulness of all moral principles, a position which, as stated above, I do not hold. [↑](#footnote-ref-60)
61. This is not to deny that this case presents any sort of dilemma. There may be a *tactical* dilemma arising out of the fact that it may be counter-productive to try to impose our view of the matter in the most ‘militant’ way possible on those directly affected by the issue. A more patient approach may be necessary. This, in my view, is the kernel of truth in the claim that cultural differences matter in cases of this sort. [↑](#footnote-ref-61)
62. It is true that there can be evidence for non-fundamental normative statements, since such statements require mind-to-world fit (see section 3.3). But the statement that a certain purpose is sufficient to justify permitting evil is a fundamental normative statement. [↑](#footnote-ref-62)