**Chapter 4: The Expressivist Challenge**

One type of antirealist who rejects both of the moral realist’s core tenets is the noncognitivist. He not only denies that there are moral facts or truths, but claims that the best way to understand moral language is nondescriptively. Why take our moral judgments, or evaluations, to refer to real moral properties existing in the world when there are no such properties? Despite the realist appearance of moral language and thought,[[1]](#footnote-1)1 noncognitivists see moral judgments as primarily evaluative in nature. The view is that moral evaluations express the speaker’s attitudes or feelings toward the object being judged; thus, we should understand the utterance, “Torturing small children is morally wrong”, as an expression of the speaker’s attitude of disapproval toward torturing small children *rather than* as an expression of the speaker’s *belief* that torturing small children has the property of being morally wrong. Moral wrongness, the noncognitivist says, is not something that we can be cognizant of in the way that we can come to know that lemons have an ovoid shape. For the noncognitivist, evaluating and describing are importantly distinct.[[2]](#footnote-2)2

In Chapter 1 we saw that, while certain moral phenomena seem to support a realist view of morality, other moral phenomena lead one to believe that moral realism cannot be true. Two features of morality stand out in terms of giving support to a noncognitivist view of moral judgments: first, because we are often motivated to act in accordance with our moral convictions, it seems that there is an internal, or necessary, connection between our moral convictions and motivating states. As we saw in Chapter 2, this feature of morality presents a problem for the realist who is committed to the mind-independence claim; realist moral properties end up looking “queer” relative to other realist entities if, upon perceiving a moral fact or truth, we are compelled to care about it and see it as giving us a reason for acting. The second feature is the existence of widespread disagreement regarding what is morally right and wrong. The apparent pervasiveness of moral disagreements undercuts the view that there is any objective support to be found for our moral value judgments. There is some initial plausibility therefore in thinking that what people are doing when expressing their moral opinions is just that—expressing their opinions, oftentimes with the aim of pressuring others to do something, or to take a similar attitude toward some “object” in the world.

At present, expressivism looks to be the most promising form of moral noncognitivism. So if it runs into difficulty, we can take it that other, less-sophisticated forms of noncognitivism will too. Simon Blackburn and Allan Gibbard have been two of the most prominant exponents of expressivism. In this chapter I shall focus solely on Blackburn’s arguments defending the thesis against one especially acute problem, since he has given that problem the most attention, and since I have seen nothing in the literature to make me think that Gibbard’s slightly different form of expressivism is immune to it.

Blackburn’s expressivist aims to be a *quasi-realist* rather than an error theorist. The expressivist who is a quasi-realist attempts to show that “the apparently ‘realist’ appearance of ordinary moral thought [and discourse] is explicable and justifiable on an anti-realist picture” (Blackburn 1993a: 151). The quasi-realist expressivist, in other words, claims to earn for us the right to think and talk about morality and use moral language as we already do even though there are, according to the thesis, no moral facts or truths. If this quasi-realist project succeeds, we needn’t give up our beliefs that moral judgments can be true or false, that some people are more morally perceptive than others, that we can be mistaken in our moral judgments, that the correctness or incorrectness of moral judgments is not a function of people’s beliefs or affective responses toward the objects of evaluation, that moral disagreements can, and ought to be, resolved through rational discourse, and so forth. Nor must we give up expressing moral judgments in assertoric form, using evaluative predicates in indirect contexts, or making the kinds of moral inferences that we do. Moreover, if the expressivist can justify these practices, he avoids having to give a story about how it is that we have fallen into such great error.[[3]](#footnote-3)3

Our concern in this chapter is whether the quasi-realist expressivist[[4]](#footnote-4)4 can offer a better account of common features of our moral beliefs and practices than the realist can. If this expressivist can get us everything the moral realist can but without the realist’s additional metaphysical commitments, then moral realism, from the standpoint of the explanatory criterion of reality, is seriously threatened. In what follows, however, I argue that expressivism fails to account for a very important feature of our moral practice. At the end of the chapter I point out that it runs into a number of other problems as well. The upshot is that expressivism is *not* explanatorily adequate, and hence, does not give us reason enough to dismiss moral realism as a candidate in our search for the kind of objectivity that we can justifiably attribute to our moral value judgments.

*The Role of Stroud’s Strategies*

Before looking at the problems facing the expressivist it will be helpful to the reader, I think, if I say a little about how Stroud’s strategies come into play in the arguments that follow. In Chapter 3 we looked carefully at the details of Stroud’s arguments with the hope that they would shed some valuable light on the debate between the moral realist and his antirealist opponents. Stroud’s strategies do bring something new to the debate, but we won’t really see this until Chapter 5. The main objection against the expressivist that is discussed in this chapter (the Frege-Geach objection) has been around for some time. It turns out that the strategy it involves is the very same as that which Stroud relies on most when arguing against the sensation account of colours. But it is exactly for this reason that Stroud’s work doesn’t shed any significant *new* light on it. Stroud’s arguments will be far more helpful, I think, to the issues raised in the penultimate section of this chapter (especially problems 2, 3, 4, and 5). I don’t try to show how they might be of help, though, because for our purposes it will be enough to address the Frege-Geach objection.[[5]](#footnote-5)5

It is important to address the latter problem, despite all the attention it has received, for the following reasons: First, Blackburn has continued to insist that he has given an adequate response to the objection (1998: section 3.4). Second, Bob Hale, who has perhaps done the most to shore up the expressivist’s position, is, I think, overly optimistic regarding the prospects of the expressivist who is a *modest* quasi-realist (1993a). Third, the Frege-Geach objection is something the expressivist *has* to address; he cannot bite the bullet on this one and say that in spite of this difficulty he has, in the end, the better account; if he isn’t able to satisfactorily address the problem, we can conclude that expressivism is untenable (and so too, in all likelihood, are all other forms of noncognitivism).

Because Stroud’s strategies are not explicitly put to use in the main part of this chapter, we should at least note here, if only very briefly, why there is reason to think that they can shed additional light on the moral noncognitivist’s prospects.

Because it is a form of noncognitivism, expressivism is susceptible to many of the types of problems facing the sensation account of colours. Here are five of those problems:

(a) The sensationalist doesn’t seem to have a way to distinguish between sensations of different colours because he doesn’t seem to have a way to specify any determinate content for them (Stroud 2000: 99-101). If we cannot distinguish a “sensation” of yellow from a “sensation” of green, how will we be able to properly attribute these sensations to oneself or to others? What would it mean to *think* of ourselves as having such sensations?

The moral noncognitivist faces the same problem. What distinguishes moral approval (or disapproval) from other kinds of approval (or disapproval)? How do we distinguish the typical moral response to cruelty from the typical moral response to rudeness? Because this problem is so significant and is also encountered by the dispositionalist and the error theorist (whether for moral properties or colour properties), I take it up in Chapter 5.

(b) The sensationalist has to explain why we shouldn’t also have a sensation account of shapes and other perceivable properties; is it not true that oranges are round and lemons ovoid? Because “[i]t seems undeniable that we do intelligibly ascribe certain properties to objects that are not “sensations” and that we also sometimes perceive objects to have those very properties” (Stroud 2000: 114), the sensation account of colours is discontinuous with our understanding of other properties that we can perceive and ascribe to objects.

The expressivist’s account of moral language is similarly discontinuous with our understanding of other regions of discourse (particularly assertion) that have the very same syntactical features (unless the expressivist is willing to be a full-blown idealist). Expressivism commits us to at least two theories of meaning and, as we will see below, at least two notions of logical inconsistency.

(c) Predicational seeing is completely mysterious on the sensation account. We take ourselves to see lemons to be yellow; predicational seeing, in this case, involves seeing lemons to have a certain property—yellowness. But it doesn’t make sense to say that yellowness is the distinctive feature of a sensation of yellow that we then project on to lemons—for the same reason it doesn’t make sense to say that painfulness is a feature of pins. This was the problem that Stroud emphasized most in his discussion of the sensation account. The sensationalist isn’t able to account for the connection between colour as it figures in thoughts, beliefs, and other psychological attitudes and colour as it figures in perception. While he can give some account of the first sample sentence of SS-I (of Chapter 3), he isn’t able to account for the way colour terms are used in the other sentences of SS-I.

This inability is exactly analogous to the kind of problem the expressivist has handling evaluative predicates in indirect contexts, the problem that the Frege-Geach objection raises. It doesn’t make sense to interpret evaluative predicates expressively when they are found in indirect contexts, at least not in the same way they are interpreted expressively when found in direct contexts. Accordingly, the expressivist has trouble making sense of the range of sentences using moral terms that we attribute to others. This is the same problem that we also find in issue 2. below (at the end of the chapter).

(d) The sensationalist also has trouble explaining nonveridical cases. It is still true to say that Jones sees a yellow lemon even if she neither sees yellow nor sees a lemon nor sees the lemon to be yellow.

As far as I can tell, the moral noncognitivist doesn’t face this type of problem.

(e) Sensation accounts, whether for pains or colours, still need to explain why it is that we have the sensations we do when the circumstances under which these sensations are usually had are met. For instance, when a pin pricks us, why do we experience a sensation of pain rather than a perception of yellow or a perception of something cubical (Stroud 2000: 98)? Why do we find the correlations we in fact find between various psychological and physical phenomena (ibid., 91-93)?

Similarly, the expressivist has to explain why we have settled upon the feature-response pairs that we have (disapproval of the gratuitous infliction of pain, for example); why are they what they are and not something else? Blackburn and Gibbard try to explain these pairs from an evolutionary standpoint; perhaps some of the feature-response pairs involved in moral practice arose because they happened to be conducive to survival. Even so, these feature-response pairs still need to be justified in moral terms. (See issue 3. below regarding the incongruity between noncognitivism and the supervenience claim.)

In general, Stroud helps us to see that the expressivist needs to make sense of moral language and thought just as the colour antirealist needs to make sense of the differences and interconnections between the sample sentences of SS-I. He also helps us to see that, because there are no direct tests for the existence or non-existence of moral properties, the only “data” we have to work with are our thoughts and beliefs (etc.), and the thoughts, beliefs, perceptions, and other psychological attitudes involving moral predicates that we attribute to others. Stroud thus helps us to see why our quest becomes one of answering the question: Can we understand these beliefs, thoughts, etc. without having to think that they represent anything that is part of the world independently of us? Finally, despite the many differences between moral properties and colour properties, I think that Stroud’s arguments against the sensation account, at least those of (a), (b), (c), and (e), lend some weight to the parallel arguments made against the expressivist, since none of these four problems hinges on the particulars of the properties themselves. Therefore, if we are persuaded that the (a), (b), (c), and (e) arguments effectively undermine the sensation account, we shouldn’t be surprised to find that the parallel arguments effectively undermine expressivism. In fact, we should expect expressivism to be susceptible to these arguments exactly because they are effective against the sensation account and do not depend on the peculiarities of the properties involved.

*The Frege-Geach Objection*

It is part of our moral practice to work out the implications of our moral views; we draw inferences from the moral judgments that we make, and we rely on both evaluative and non-evaluative premises in our moral reasoning. In doing so, we assume that evaluative judgments can enter into logical relations with other statements in the same ways that non-evaluative judgments can. For example, if we accept

P1 : If it is wrong to tell lies, it is wrong to get your little brother to tell lies.

and

P2 : It is wrong to tell lies.

we are likely to infer

C : It is wrong to get your little brother to tell lies.

The inference is valid, we say, by application of modus ponens; we say that C *logically follows* from P1 and P2 (or equivalently, ‘P1 and P2 logically entail C’) because we understand P1 to have the form ‘*P*→*Q*’ and P2 to have the form ‘*P*’. This understanding of what licences the above inference, however, seems to commit us to a realist interpretation of the sentences.[[6]](#footnote-6)6 For in order to see the inference as a typical application or instance of modus ponens, we must interpret the sentences as having truth values. The notion of validity that we take ourselves to be relying on also requires that the sentences be truth-evaluable: we say that the inference is valid just in case it is not possible for all the premises to be true and the conclusion false. We say, too, that P1, P2, and not-C are *logically inconsistent*.

The expressivist, if he doesn’t want to be an error theorist, must agree that we are justified in taking the above argument to be an instance of valid reasoning, and that we are justified in thinking that the mistake one makes in denying C but affirming P1 and P2 is a kind of *logical* mistake. The challenge for this expressivist is to explain why we are justified in thinking this. The path of least resistance, it seems, would be to show how it is that the above argument is in fact an instance of valid reasoning, and to show how it is that someone who denies the conclusion but affirms the two premises is in fact making a logical mistake. Still, this will be a difficult task for the expressivist since he holds that C and P2 lack truth values; hence he must resort to notions of ‘logically valid’, ‘logically inconsistent’, and ‘logical entailment’ that are different than the ones we ordinarily employ. Moreover, it seems that he won’t be able to explain the validity of the above argument in terms of modus ponens. This follows because the expressivist holds that to affirm P2 is to express a negative attitude toward telling lies. Yet since affirming P1 doesn’t require one to hold this attitude, it is hard to see how the expressivist, or any other kind of noncognitivist, can construe the above argument as having the form, ‘if *P*→*Q* and *P*, then *Q*’.

It is because the compound sentences of ordinary propositional logic are truth-functional that we can apply rules like modus ponens. Since the expressivist denies that P1 is truth-functional and that P2 and C have truth values, he has to find other ways in which P1 and P2 are logically related if he is to explain why we are justified in inferring C from them. Yet this looks to be an impossible task if the expressivist is committed (as it seems he must be) to the view that the semantic role of evaluative predicates radically changes depending on the contexts in which these predicates appear; the expressivist is committed to saying that ‘is wrong’ in P2 and C has an expressive role, but in P1 the predicate—because it is embedded in a conditional—cannot function expressively in exactly the same way.[[7]](#footnote-7)7

In the early 1960s Peter Geach argued that this result is also highly problematic for noncognitivist views because it violates “the Frege Point”—the idea that the meaning of a sentence or proposition remains constant across contexts.[[8]](#footnote-8)8 Without this constancy of meaning any modus ponens inference that involves sentences with the same evaluative predicate found in different contexts will be invalid.[[9]](#footnote-9)9 If the meaning or content of ‘It is wrong to tell lies’ changes in any significant way from P1 to P2, we commit the fallacy of equivocation if we infer C.

Simon Blackburn has attempted on a number of occasions to show how the expressivist might meet this challenge. He has offered at least three distinct responses; the first appears in the last section of his paper, “Moral Realism” (1971), which I shall abbreviate as MR. The second is found in Chapter 6 of his book, *Spreading the Word* (1984; hereafter STW); this second response is an extension and refinement of the first. A third solution, appearing in the paper, “Attitudes and Contents” (1988; hereafter AC), was necessary due to forceful criticisms of the previous two by G. F. Schueler and others.1[[10]](#footnote-10)0 Blackburn also discusses this third solution in section 3.4 of his book, *Ruling Passions* (1998). Bob Hale offers a very detailed critique of it in his paper, “Can There Be a Logic of Attitudes?” (1993a). Hale shows why the third solution, as it stands, is inadequate. In “Expressivism and Irrationality” (1996), Mark van Roojen presents further reasons for concluding that the expressivist’s story is explanatorily weak.

In what follows I look at each of Blackburn’s three solutions and the problems facing them. Like van Roojen, I think it highly unlikely that the expressivist can come up with a satisfactory solution to the problem of indirect contexts. Even if he should find a way to get around the difficulty that Hale points to, the solution he finds will almost certainly have to be subtle and elaborate; there is reason to worry, then, that we won’t see it as an equal or better explanation than the straightforward one that a moral cognitivist can offer. (The moral cognitivist can explain our practice of “making inferences based upon and leading to evaluative judgments”1[[11]](#footnote-11)1 simply by reiterating our current understanding of what we are doing).

In MR Blackburn proposes that we view the conditional, P1, as an *assertion*: the speaker is claiming that disapproval of telling lies involves disapproval of getting little brother to tell lies (1993a: 127). At the same time, we should view P2, not as a factual or descriptive statement, but as expressing an attitude. Blackburn then claims that it still makes sense under this interpretation to say that it is logically inconsistent to deny C but affirm P1 and P2. He offers two reasons in support of this claim. First, in asserting P1 the speaker is expressing a moral standard or commitment (127); the idea seems to be that the speaker who sincerely utters P1 must (in some sense of this word) hold an attitude of disapproval toward getting little brother to tell lies if she disapproves of telling lies; otherwise she is not sincere in uttering P1. Second, Blackburn claims that denying C after affirming P1 and P2 would be a violation of modus ponens.

We should see immediately, however, that modus ponens—as we ordinarily understand it—is not preserved under this expressivist interpretation of the argument. P2 expresses an attitude, whereas P1 is a claim about what a particular attitude involves; as such, its antecedent never expresses an attitude at all. Van Roojen writes: “Our standard accounts of the validity of modus ponens require that P2 express what the antecedent of P1 contributes to what P1 expresses. On the noncognitivist account, the words of the antecedent of P1 and the words of the second premise, P2, are functioning as homonyms” (1996: 313). The premises therefore lack the logical form needed for an application of modus ponens.1[[12]](#footnote-12)2

How much of a problem is it for the expressivist if he cannot interpret the above argument as an instance of modus ponens? As Schueler observes, it is a bit of a set-back because the quasi-realist’s aim is to explain the realist appearance of ordinary moral thought (1988: 495), and the type of inference in question seems to be a very prevalent feature of our moral practice. We say that the above argument is valid because it is an instance of modus ponens; we *make* the inference because we see it as being justified by modus ponens. The expressivist who is also a quasi-realist should want to make sense of this understanding because the understanding itself is part of our moral practice. (It won’t do for him to reply that the quasi-realist project is limited to explaining only the realist appearance of *first-order* moral practices, that he need not explain any realist understanding of those practices.1[[13]](#footnote-13)3) Nonetheless, if the expressivist cannot succeed with this part of the quasi-realist project, he could try to put a good face on things and say that what really matters is having an account of why the inference is valid, and an account of why a person is making a logical mistake if they affirm P1 and P2 but deny C. Afterall, the expressivist’s story, in the end, only needs to be stronger than competing accounts, and perhaps there are problems of an even more serious nature facing the competing accounts.

What about Blackburn’s first reason, then, in support of his claim that we can take P1 to be an assertion, P2 and C to be expressions of attitude, and still say that it is logically inconsistent to affirm the premises and “deny”1[[14]](#footnote-14)4 the conclusion? If Blackburn’s other argument in support of this claim goes through, the expressivist is clear of the two most critical problems that the Frege-Geach objection raises for him. He can say that the example argument is valid in the sense that affirming the premises and denying the conclusion would be logically inconsistent; and he can say that the person who denies the conclusion while affirming the premises is making a logical mistake. Schueler, though, has forcefully shown why Blackburn’s first argument also fails. Whereas there may be reason to say that a person is inconsistent if they believe P1, hold a negative attitude toward telling lies, and yet tolerate getting little brother to tell lies (one might, for example, try to equate inconsistency with lack of coherence among mental states), there is no reason, as yet, to see the inconsistency as being of a logical sort.1[[15]](#footnote-15)5 Thus there is no reason, as yet, to see the argument as valid in the way just mentioned, nor any reason, as yet, to say that this person is making a *logical* mistake.

Schueler recognizes (1988: 499) that this last criticism may be a bit unfair to the expressivist, for he will almost certainly have a different idea of what counts as a logical mistake.1[[16]](#footnote-16)6 But until such an account is given, he cannot legitimately say that he has a solution to the Frege-Geach objection. Moreover, the expressivist still has to explain why evaluative predicates demand such different treatments depending on whether they are found in asserted or unasserted contexts. In the unasserted contexts of P1, the predicates refer to attitudes; in the asserted context of P2, the predicate expresses an attitude.

Let’s move on to Blackburn’s second response, as I don’t see how the MR treatment of evaluative predicates in indirect contexts could be made to work as the expressivist needs it to. The value of considering each of Blackburn’s solutions is that we get a clear sense of the difficulties that the Frege-Geach objection poses for the expressivist. This, in turn, should help to persuade us that the issue is a genuine problem for him.

In STW Blackburn continues to say that we should see P1 as expressing a moral point of view (1984: 192-93), but that we should no longer see the conditional as an assertion. Instead, we should interpret P1 as expressing the speaker’s *attitude* toward joining the attitudes (or beliefs) that would be expressed if the components were found in direct contexts. We should thus see P1 as expressing a favorable higher-order attitude toward combining disapproval of telling lies with disapproval of getting little brother to tell lies.

Blackburn develops this idea further by showing how we might construct an expressive language. He asks us to

Imagine a language unlike English in containing no evaluative predicates. It wears the expressive nature of value judgments on its sleeve. Call it Eex. It might contain a ‘hooray!’ operator and a ‘boo!’ operator (H!, B!) which attach to descriptions of things to result in expressions of attitude. H!(the playing of Tottenham Hotspur) would express the attitude towards the playing. B!(lying) would express the contrary attitude towards lying, and so on. (1984: 193)

We can denote attitudes themselves with the use of bars: ‘|H!(*X*)|’ refers to approval of (*X*). And we can “use the semi-colon to denote the view that one attitude or belief involves or is coupled with another” (1984: 194). The argument we have been discussing will then look as follows:

P1' : H!(|B! (lying)| ; |B! (getting little brother to lie)|)

P2' : B! (lying)

C' : B! (getting little brother to lie)

Blackburn tells us that

Anyone holding [P1' and P2'] must hold the consequential disapproval: he is committed to disapproving of getting little brother to lie, for if he does not his *attitudes clash*. He has a fractured sensibility which cannot itself be an object of approval. The ‘cannot’ here follows not (as a realist explanation would have it) because such a sensibility must be out of line with the moral facts it is trying to describe, but because such a sensibility *cannot fulfil the practical purposes for which we evaluate things*. Eex will want to signal this. It will want a way of expressing the thought that it is a logical mistake that is made, if someone holds the first two commitments, and not the commitment to disapproval of getting your little brother to lie. (1984: 195; my emphasis)

As Schueler notes (1988: 498), what is interesting about the STW response is that Blackburn never claims to have shown that the inference to C (or C') is valid; nor does he claim to have shown that the inference is an instance of modus ponens.1[[17]](#footnote-17)7 What Blackburn does claim to have shown is that we are involved in a logical mistake if we accept P1 and P2 but then deny C.

But should we accept even this claim? In expressing P1 the speaker, Blackburn tells us, is expressing a moral point of view. So if any kind of mistake is involved in “denying” C but affirming P1 and P2, it would seem to be a moral, rather than a logical, one. (We will return to this point below.) Furthermore, unless we make some important changes to the semantics of Eex, it isn’t the case that there is always, or even ever, a clash of attitudes if we accept premises like P1 and P2 but deny a conclusion like C. Schueler illustrates this point as follows:

Suppose I have the following attitudes: I disapprove of lying. I approve of the view that disapproval of lying “involves” disapproving of getting one’s little brother to lie. And (to make it simple) I approve of getting my little brother to lie. Here are three different attitudes. What is supposed to make us think that they somehow “clash”? We cannot say (or anyway Blackburn cannot say) that disapproval of lying “clashes” with approval of getting my little brother to lie, period. Even though there is a sense of ‘disapprove’ in which that seems true, it is not a point Blackburn can use since appeal to it would simply circumvent the hypothetical premise in this argument and, clearly, that has to do at least some of the work if this whole account of modus ponens isn’t to just fall apart. (1988: 499)

Keep in mind that if there is no clash of attitudes, it is not clear that *any kind* of mistake is being made. Hale identifies the problem with Eex that leads to this difficulty:

[I]t is just not clear that approval of making one attitude ‘follow upon’ another1[[18]](#footnote-18)8 excludes toleration of holding the second attitude without subscribing to the first. The gap is, perhaps, most easily perceived if we consider simple attitudes: I may surely approve of giving money to WoW without thinking it intolerable not to do so. Because of this, it does not appear that approval of *x*-ing and intolerance of not-*x*-ing are related, as Blackburn’s claim seems to require them to be . . . . But if not, there does not have to be any clash of attitudes in one who endorses [P1 and P2] but not [C]. (1993a: 343)

Hale shows how the expressivist can change the semantics of Eex to ensure a clash. One way would be to stipulate “that H!*x* is to be understood as expressing an attitude of *insistence* upon *x*” (344). Another way would be to introduce a toleration operator, T!*x*, as primitive and define the B! and H! operators in terms of it; that is, define H!*x* “as expressing refusal to tolerate failure to *x* (that is, H!*x* =df ¬T!¬*x*)” and define B!*x* “as refusal to tolerate *x* (that is, B!*x* =df ¬T!*x*)” (344). Notice, however, that whatever solution we go with, our interpretation of what we are saying and doing when making statements like P1 and P2 is becoming more distant from our ordinary understanding of what we are saying and doing. When we approve of *x*-ing, our mental state is not usually one which equates with refusal to tolerate failure to *x*; nor is it usually one which equates with insistence upon *x*. Our ordinary understanding of what we are saying and doing makes room for the gap that Hale speaks of. In any case, if the expressivist’s story gets too subtle and elaborate, he risks the danger of not having a best explanation, and he risks the danger of not being able to make sense of our current understanding of what we are doing.1[[19]](#footnote-19)9

But even if the expressivist alters Eex to ensure a clash among attitudes like P1, P2, and not-C, it is still not clear whether any kind of mistake is involved in holding on to attitudes that clash. In what way is it a mistake to hold directly conflicting attitudes? How am I mistaken if I both approve of lying and disapprove of lying? Blackburn implies that we are *practically irrational* if we endorse attitudes that clash (one “has a fractured sensibility which cannot itself be an object of approval”; “such a sensibility cannot fulfil the practical purposes for which we evaluate things” (1984: 195)). But are we? It seems that there is more than one point of view from which we might say that it may very well be to an agent’s advantage to approve of (or tolerate) lying at time *t* and in circumstances C and yet disapprove of (or not tolerate) lying at time *t*2 and in circumstances C2. It is easy to think of other examples. Blackburn’s response doesn’t tell us, for instance, how it is problematic for an agent, from the standpoint of practical rationality, to approve of courage when it is required of others but disapprove of it when it is required of the agent himself, or of his children.2[[20]](#footnote-20)0 Schueler notes, in addition, that we can have “clashes” of attitude

without being “mistaken” about anything (or otherwise having [our] practical purposes thwarted by “fractured sensibilities”). Almost every time I walk past the cookie jar . . . I experience such a conflict of attitudes. On the one hand, I approve of my eating a cookie since, from long experience, I know I will enjoy it. On the other hand, I also disapprove of eating one since, from equally long experience, I know what it does to my waistline. So here is a clash of attitudes, but where is the “mistake” . . . ? The fact that I cannot act on or satisfy both attitudes just means that I have to figure out what to do here, not that there is anything untoward or fishy about having both attitudes. (1988: 500)

Perhaps the expressivist will reply by saying that we shouldn’t have to figure out what to do when we walk past the cookie jar, since we can’t be very effective agents if we are always caught up in conflicts of this kind. Even so, it is still not clear from what Blackburn says that this type of problem amounts to making some kind of mistake.2[[21]](#footnote-21)1

For the sake of argument, let’s say that Blackburn *has* shown that there is a kind of mistake in having a clash of attitudes. (As we have just seen, the expressivist who makes use of Hale’s refinements to Eex does at least have a way of showing that there is some sort of clash of attitudes in the person who accepts P1 and P2 but denies C.) Will the expressivist have then shown what he needs to show? No, for he still hasn’t shown that endorsing attitudes that clash in the way described amounts to a kind of logical mistake. We can at most say that the mistake is of a moral sort. The fault “consist[s] in holding a combination of attitudes [in this case, P2 and some form of not-C] of which you expressly disapprove [when you endorse P1]” (Hale 1993a: 344). Therefore, the STW response, like the MR response, doesn’t convincingly show that we are involved in a logical mistake if we accept the premises but deny the conclusion. The STW response thus fails one of its two critical tests.

What about the other critical test, that of the inference, if P1 and P2 then C, being valid? The concern with validity, ordinarily, is that arguments be truth-preserving. What the expressivist might be concerned with is that “arguments” preserve consistency. He could then try to employ something like the following notion of validity: An argument is valid if and only if the premises entail the conclusion, that is, if and only if the set of sentences consisting of the premises and the negation of the conclusion is inconsistent. In fact, if the expressivist relies on this notion of validity and also provides us with a notion of *logical* inconsistency, he may be able to meet both critical tests (i.e., he may have a response to both of the two most critical problems that the Frege-Geach objection raises for him). As we have just seen, however, the STW response isn’t up to this task, for that response doesn’t make it at all clear how the person who accepts P1 and P2 but rejects C is involved in a logical mistake. Let’s move on then to Blackburn’s third response, the solution found in AC. The virtue of the third solution is that it does provide us with a notion of logical inconsistency.

The third response is fairly complicated, but we need to delve into some of its details if we are to get a sense of the lengths to which the expressivist has gone, and the lengths to which he has yet to go if he is to meet the challenge of indirect contexts. Fortunately Bob Hale (1993a) gives an excellent summary and critique of it. Hale writes:

In matters of syntax, at least, the theory of AC diverges sharply from that of STW. In Eex, the operators H! and B! are applied to *descriptions*—gerundival phrases like ‘lying’, ‘getting others to lie’, etc.—and so are not, properly speaking, iterable. In the new theory, H! is retained, and the toleration operator T! is added. But now they behave as regular sentence-forming operators on *sentences*. Compounds so formed, such as H!*p*, T!(*p* & *q*), and the like, are allowed all the sentential embeddings that plain *p* can undergo. The syntax thus provides for iteration and for sentences expressing higher-order attitudes, such as H!(B!*p* → B!*q*) corresponding, roughly, to the earlier form H!(|B!*x*| ; |B!*y*|) . . . .

We have, then, in place of Eex, something that looks very much like the formal languages of familiar intensional extensions of the language of standard first-order logic, with attitude operators supplanting the more familiar intensional operators. The thought is that we can adapt techniques already to hand for, for example, deontic logic to produce a generalized logic of attitudes which does justice, *inter alia*, to our sense that one who refuses to accept inferences such as [if P1 and P2, then C] is illogical, not just depraved. (1993a: 344-45)

Blackburn modifies a deontic logic developed by Hintakka.2[[22]](#footnote-22)2 “In Hintakka’s semantics, the central notion is that of norms obtaining in a possible world, and of the deontic alternatives to that world being the possible worlds that are in accordance with those norms” (Blackburn 1993a: 193). If we substitute ‘attitudes’ for ‘norms’ in the previous description, we get a very rough sense of what Blackburn has in mind. The logic is such that, if *A* is a well-formed formula in the language we are working with, then “Logical truth of *A* is unsatisfiability of the negation of *A*; *B* is a logical consequence of *A* if and only if *A*→*B* is valid, that is (*A* & ~*B*) is unsatisfiable” (ibid.). Hintakka’s logic is then adapted as follows:

Suppose that to a standard first-order language we add operators H! and T! subject to the condition that if *A* is a well-formed formula, H!*A* is well formed and T!*A* is well formed. Suppose now we start with a set of sentences *L*, which may contain sentences with these operators among them. We begin by defining a *next approximation to the ideal*, *L*\*, of *L*.

(1) If H!*A* ∈ *L*, then H!*A* ∈ *L*\*;

(2) If H!*A* ∈ *L*, then *A* ∈ *L*\*;

(3) If T!*A* ∈ *L*, then a set *L*\* containing *A* is to be added to the set of next approximations for *L*;

(4) If *L*\* is a next approximation to the ideal relative to some set of sentences *L*, then, if *A* ∈ *L*\*, *A* ∈ subsequent approximations to the ideal *L*\*\*, *L*\*\*\*, . . . .

(Blackburn 1993a: 194)

Blackburn now understands the attitudes that might be involved in our moral responses in terms of *aims or goals*: “H!*p* expresses the attitude of endorsing the goal *p*” (ibid., 192).2[[23]](#footnote-23)3 So, in Hale’s words: “Rules (1) and (2) stipulate, in effect, that whilst a goal may go unrealized in the actual world (that is we may have H!*A* but ¬*A*), it must be fulfilled in all approximations to the ideal (relative to our initial attitude set), and must remain in force as a goal in those approximations. Rule (3) sees realization of a toleration as compatible with perfection, but not required for it” (1993a: 346). Very roughly, rule (3) simply requires that one of the next approximations contain *A* if T!*A* ∈ *L*.

The aim of Rule (4) is to ensure that any *A* that makes it into an approximate ideal also makes it into all subsequent ideal possible worlds, unless the *A* in question is derived from the realization of a toleration. However, Hale shows that Rule (4) leads to some unwanted results and fails to get us some of the results that we do want. He thinks these problems can be avoided if we replace (4) with “(4)' : For any *L*\* from *L*, if H!*A* ∈ *L*\*, then *A* ∈ *L*\*” (1993a: 349).

It is worth pausing for a moment to remind ourselves what the expressivist wants to accomplish with this technical apparatus. He wants to show how an expressivistic interpretation of evaluative predicates still permits us to say that an inference like the one from P1 and P2 to C is valid, and that we are justified in thinking that the person who accepts P1 and P2 but rejects C is making some kind of logical mistake. The expressivist would also like to be able to say why we think this particular inference is an instance of modus ponens.2[[24]](#footnote-24)4 We saw that the expressivist has to rely on a notion of validity which doesn’t presuppose that the sentences reflecting the inferences we are making are truth-apt. The AC proposal is to base validity on a notion of inconsistency that is in turn understood in terms of unsatisfiability. Recall that “*B* is a logical consequence of *A* if and only if *A*→*B* is valid, that is (*A* & ~*B*) is unsatisfiable” (Blackburn 1993a: 193). If, under the expressivist interpretation offered in AC, we can show that P1 and P2 logically entail C, we then can legitimately say that the person who accepts P1 and P2 but denies C is making a kind of logical mistake; for they will be holding a set of attitudes (aims or goals) that is not realizable in any possible world (given rules (1) - (4)' of the deontic logic).

Regarding the above technical apparatus, Hale writes:

The intention is that when these rules [i.e., (1)-(4)], together with semantic tableau rules for ordinary sentential operators, are applied to an initial set of sentences L, they will generate a partially ordered set of sets of sentences (next ideals). An ideal is said to be *final* when further applications of the rules produces no new sentences. If L, or any next ideal, contains ordinary branching compounds, then there will be more than one *route* to a set of final ideals. What matters for the consistency of L is that there should be at least one set of final ideals providing for the realization, separately, of each of the T!*A* in or implied by L, compatible with realization of all the H!*B* in or implied by L. Thus Blackburn defines:

A set of sentences L is *unsatisfiable* iff each route to a set of final ideals from L terminates in at least one set of sentences containing both *A* and ¬*A*, for some *A*.2[[25]](#footnote-25)5

(Hale 1993a: 346)

Let’s apply the logic to the argument we have been working with since the beginning of this section. Let *p* stand for “It is wrong to tell lies” and let *q* stand for “It is wrong to get your little brother to tell lies”. Blackburn is saying that the above logical apparatus allows us to represent the argument as follows:

P1'' : B!*p* → B!*q*

P2'' : B!*p*

C'' : B!*q*

The premises entail the conclusion if it is the case that {~B!*p* ∨ B!*q*, B!*p*, ~B!*q*} is unsatisfiable. (Blackburn will want us to translate the conditional statement into a disjunction.) By the ordinary tableau rules for sentential logic, this set of sentences will be unsatisfiable if both {~B!*p*, B!*p*, ~B!*q*} and {B!*q*, B!*p*, ~B!*q*} are unsatisfiable. Since Blackburn doesn’t give us any rule for the B! operator, we either have to add a rule for it or reduce formulae with this operator to formulae involving only the T! operator, or only the H! operator. Let’s say that B!*A* is equivalent to ~T!*A* and ~B!*A* is equivalent to T!*A*. The first set then becomes {T!*p*, ~T!*p*, T!*q*} and the second set becomes {~T!*q*, ~T!*p*, T!*q*}. Before we can apply any of rules (1)-(4)' we have to get rid of the “external” negation signs. The first set becomes {T!*p*, H!~*p*, T!*q*} and the second set becomes {H!~*q*, H!~*p*, T!*q*}. Applying rules (2) and (3), we see that a next ideal of the first set contains both *p* and ~*p*, and a next ideal of the latter set contains both *q* and ~*q*. Thus, since the premises in conjunction with the negation of the conclusion is unsatisfiable, we can say that the premises logically entail the conclusion. The inference from P1'' and P2'' to C'' also appears to be an instance of modus ponens.

As far as the logic’s ability to handle our one example argument, then, everything seems good. But as Hale points out (1993a: 346), Blackburn’s third solution still runs into problems with indirect contexts, for *we have to be able to interpret the logic expressivistically*. And it is not clear how this can be suitably done when expressive compounds like H!*p* are embedded, that is, when they fall under the scope of ordinary sentential operators. We noted already that “H!*p* expresses the attitude of endorsing the goal *p*”. And we just saw that Blackburn handles the negation operator by moving it inside: ~H!*p*, he says, expresses the attitude of tolerating ~*p* (1993a: 192).2[[26]](#footnote-26)6 We are not told what attitude B!*q* expresses (we are told only that it amounts to “rul[ing] *q* out of any perfect world” (1993a: 189)), but I take it that it probably expresses the attitude of rejecting, or not tolerating, the goal *q*; or it expresses the desire or aim to see that *q* not happen. How, though, are we to interpret sentences like *p* → H!*q* or T!H!*p*? Blackburn’s logic has us treat ‘→’ in P1'' above as an ordinary sentential operator. But ordinarily, such connectives operate on terms that function descriptively, not expressively.

Blackburn’s suggestion is that we think of *p* ∨ H!*q* as “the commitment of one who is . . . *tied to a tree*—that is, tied to (either accepting that *p*, or endorsing *q*), where the parentheses show that this is not the same as (being tied to accepting *p*) or (being tied to endorsing *q*). Rather, the commitment is to accepting the one branch should the other prove untenable” (1993a: 192). Hale writes: “The proposal, as I understand it, is that we can characterize the sense of compounds whose components may be evaluative, not in terms of their truth conditions, but in terms of the inferential commitment involved in endorsing them” (1993a: 347). Hale’s understanding seems right; in *Ruling Passions* Blackburn tells us that “The key idea here is one of a functional structure2[[27]](#footnote-27)7 of commitments that is isomorphic with or mirrored by the propositional structure that we [ordinarily] use to express them” (71); the expressivist holds that our moral attitudes can and do appear in propositional clothing (it is the quasi-realist who reveals this fact). We can see what the implications of an attitude are, or a combination of attitudes, or attitudes and beliefs, by seeing what the practical effects are of holding the attitude or the combination.

However, Blackburn’s interpretative strategy doesn’t resolve *the dilemma* that he is up against (see (Hale 1993a: 339; 352-53)). Blackburn’s aim is to give an expressive interpretation to each of P1, P2, and C. But if he does, it doesn’t seem that he can show there to be any logical connection between P1 and P2, and thus it doesn’t seem that he can show the inference to be valid. For when P1 is interpreted expressivistically, the logical connective is no longer dominant, and hence, P1 cannot be treated as an ordinary conditional. On the other hand, if the expressivist interprets P1 such that the connective *is* dominant, P2 (i.e., the expression it stands for) seems to have to take on different semantic roles depending on whether it is found in direct or indirect contexts, and the interpretation of P1 is no longer truly expressivistic.

Blackburn’s interpretation doesn’t provide an escape from this dilemma because the deontic logic he employs requires that we be able to break terms apart when they occur in disjunctions; the logic has us handle ‘→’ as we do the truth-functional conditional. But his expressivist interpretation of ‘→’ has us “tied to a tree”; we are to think of *p* ∨ H!*q* as a *single* commitment, not two separate commitments joined by an ‘or’. Thus, the way the deontic logic treats this expression is in direct conflict with the expressivistic interpretation. When we are “tied to a tree”, the commitment—as Hale puts it—does not distribute across the disjunction: “We have a single disjunctive commitment, not a disjunction of commitments. Someone who affirms *p* ∨ H!*q* . . . need not be committed to affirming *p* or to endorsing *q*” (1993a: 353). Like the first two responses to the problem of indirect contexts, then, Blackburn’s third response also fails to show how our original argument is an instance of modus ponens.

How does the third response do on the question of validity, though? Hale shows how the expressivist can handle the matter of validity by employing the above deontic logic, but he will be required to make changes to how we interpret compounds with evaluative constituents. An expressivistic interpretation of such compounds requires that we interpret them with dominant attitude operators; the sentential connectives, in other words, need to be under the scope of an H! or a T!. So Hale suggests that we interpret P1, not as B!*p* → B!*q*, but as H!(B!*p* → B!*q*) or some equivalent (Hale 1993a: 354). We also need a translation for the formula B!*A*; Hales suggests that we equate it with H!~*A* (ibid.). With these interpretations in hand, we can then use rules (1)-(4)' to show that P1 and P2 do in fact entail C. (I will spare the reader the details.) In this way the expressivist can explain, at least for our one argument, why we intuitively think it is valid. Hale points out, however, that complications arise when we are faced with another kind of argument involving embedded evaluative predicates. This complication puts into doubt the possibility of an expressivist logic of attitudes.

The arguments in question are those containing what Hale calls “mixed” conditionals (Hale 1986: 74-75). A mixed conditional is one in which the antecedent or consequent is truth-apt in the ordinary sense while the other component of the conditional includes an evaluative predicate. For example, line (7) below is an example of such a conditional. The expressivist needs to be able to show that the following argument is also valid (Hale 1993a: 352):

(7) If Bill stole the money, then he should be punished.

(8) Bill stole the money.

(9) Therefore, he should be punished.

In AC Blackburn translates (7) as *p* → H!*q*, but we have seen that this won’t do; we need an attitude operator to dominate the logical connective. So Hale suggests that we try either H!(*p* → H!*q*) or H!(*p* → *q*). My own view is that only the first is acceptable if we want to respect the Frege Point. In any case, neither formula enables us to show that {(7), (8), not-(9)} is logically inconsistent by way of applying rules (1) - (4)' of the deontic logic (see (Hale 1993a: 354-55)).

Although Hale attempts to show how the expressivist might try to extricate himself from this difficulty, it is not at all clear that the strategy he offers will work. Even if it does work in the way that Hale outlines, the costs are likely to be very high. Indeed, if Mark van Roojen is right, we have little reason to expect that the expressivist’s notion of consistency could ever be extensionally equivalent to our ordinary notion. (The expressivist would then have to provide us with an error theory.)

Hale’s strategy is built around first understanding why it can make sense, from the expressivist’s viewpoint, to say that a person can consistently hold (7), (8), and not-(9). We should

think of the initial sets of sentences to which [the deontic] rules are applied, not as purporting to depict a collection of normative/evaluative attitudes held in certain circumstances but, as characterizing the combined moral attitudes *and beliefs* of some single (possible) subject. The consistency-question we are concerned with, in other words, is not whether, *given that the facts are thus and so*, it is consistent to adopt such and such a collection of normative attitudes: but whether, *given that the subject has such and such beliefs about how things are*, he can consistently hold those normative attitudes. (1993a: 355; Hale’s emphasis)

So in the case of our example argument, a person can consistently endorse (7) and not-(9) if what they believe, (8), is false. If in reality Bill did not steal the money, I can consistently believe, *per the expressivist’s present notion of consistency* (one which relies on the deontic logic), that Bill stole the money, if he stole the money then he should be punished, and Bill should not be punished. Of course, it is a major defect of this view that we don’t ordinarily say that these three beliefs or attitudes are consistent.

Strangely, Hale expects us to say that the person who holds these three beliefs is making a logical mistake, and that if the person herself is sufficiently reflective, she, too, should see that she is making a logical mistake.

we cannot, or so it seems to me, avoid this conclusion: that a sufficiently reflective subject who endorses (7) and (8) but rejects (9) will be able to see that his normative attitudes are jointly realizable only at the cost of his factual belief. He should recognize—should he not?—that something is amiss, and that he is logically obliged either to revise his belief or jettison one of his normative attitudes. (1993a: 356)

But Hale cannot have it both ways. The expressivist’s logic of attitudes *doesn’t* show, as yet, that this person is making any sort of logical error; it in fact shows just why it is that the person is not in error.

Hale implicitly concedes this last point when he goes on to explain why it is that his solution isn’t complete. He recognizes that the logic of attitudes that he has set out requires an extended notion of consistency, but one that needs to be constrained (357). Unfortunately, it is not at all clear how it is to be appropriately constrained. When should we say that a *combination* of beliefs and attitudes is inconsistent? We say that a set of beliefs is consistent “when it is possible for all the beliefs in the set to be true together” (Hale 1993a: 357). And we just saw, with Blackburn’s third response, that the expressivist can say that a set of evaluative attitudes is consistent “if there is a system of possible worlds (meeting certain constraints) which realizes it” (ibid.). But we get very undesirable results if we simply combine these two notions of consistency; for example, under a straightforward combination of the two notions it will be inconsistent to believe not-*p* but hold that *p* ought to be the case (ibid.). What is needed, Hale says, is a non-arbitrary “fixity operator” for a subject’s beliefs; when trying to determine whether a subject’s “total attitude set” is consistent, certain of the agent’s beliefs may need to be treated throughout the process as always true (357-58). For example, if we are to show that (7) and (8) logically entail (9) under *this* new extended notion of consistency (combine the notion of consistency for beliefs with the notion of consistency for evaluative attitudes and add a fixity operator), we will want to “fix” (8) as true. The problem is knowing when it is that a belief needs to be fixed in this way in order for the expressivist to get the kind of results that we expect. The big question we are left with at the end of Hale’s paper, then, is whether there are non-arbitrary general criteria the expressivist can employ for constraining the extended notion of consistency referred to. To my mind, this is no small hurdle that the expressivist has to overcome.

Van Roojen’s article adds to our reasons for thinking that the Frege-Geach objection poses insurmountable obstacles for the expressivist. As we just saw, the expressivist requires an extended notion of consistency, but one that needs to be constrained by a fixity operator for beliefs; without the fixity operator, sets of sentences that we ordinarily take to be inconsistent will be classified as consistent. This means, in turn, that the expressivist, as yet, has no way to explain the validity of certain arguments that we ordinarily say are valid. Van Roojen shows that the expressivist also faces pressure from the other direction. For the expressivist has also extended our ordinary notion of *in*consistency: he is now talking about attitudes being inconsistent, and combinations of attitudes and beliefs being inconsistent. Ordinarily we don’t say that one is involved in a logical contradiction if they hold that a certain state of affairs is good while also holding that the alternatives to that state of affairs are good; nor do we think it logically impossible for there to be moral reasons for doing X but equally strong moral reasons for not doing X (Van Roojen 1996: 321-22). But the expressivist, it would seem, needs it to be the case that H!*A* and ~H!*A* are unsatisfiable; likewise, he needs it to be the case that H!*A* and H!~*A* are unsatisfiable. (These constraints, anyway, are what the expressivist of AC holds to.) Van Roojen cogently argues that this expanded notion of inconsistency won’t track our ordinary notion of inconsistency. And why should we think it ever could? In broadening the notion of inconsistency one is making it “less stringent” (ibid., 330); more sets of sentences will count as inconsistent. The expanded notion of inconsistency, in other words, will find inconsistencies where our ordinary notion won’t; it will thus permit certain arguments to be valid that we wouldn’t ordinarily say are such.

This unwanted result that accompanies the expressivist’s less stringent notion of inconsistency is not the same problem as that for which the fixity operator on beliefs is needed. So the fixity operator, assuming it can be found, won’t be able to do double-duty. As van Roojen emphasizes (321), Blackburn has extended the notion of inconsistency by incorporating substantive moral assumptions into the deontic logic of AC. These moral assumptions are necessary in order for the expressivist to explain the validity of certain standard inferences. It is a substantive moral claim to say that what H!*A* represents “rules out” what H!~*A* represents, and to say that what H!*A* represents rules out what T!~*A* represents. It is a substantive moral claim to say that if there are moral reasons for doing X then there cannot be equally strong moral reasons for not doing X (or for doing something which is incompatible with doing X). It is also a substantive moral claim, or so I would argue, to say that the attitudes that are involved in our moral responses can be understood in terms of aims or goals, and to assume that moral attitudes, so-understood, can be realized (in practical terms) in just one way. We saw, too, that the AC solution equates T!*~A* with ~H!*A* and ~T!*A* with H!~*A* (Blackburn 1993a: 192). (My suspicion is that the AC solution will also have to reduce disapproval of *A* to ~T!*A*; it doesn’t seem that we will get the contradictions that we expect if we can’t define disapproval in terms of the other two primitives.)

There are at least three important things to note regarding these substantive moral assumptions. The first is a point already noted: the equivalencies just mentioned fail to respect ordinary understanding; for example, we take there to be a gap between approval of *x*-ing and not tolerating not-*x*-ing. Also, I doubt that we want to say that disapproving of agents doing *x* always means that we approve of agents not-*x*-ing—that we can always understand a con-attitude in terms of a pro-attitude, and vice-versa. This failure to respect ordinary understanding is reason enough to think that the expressivist’s extended notion of inconsistency—even after additional tweaking—won’t be extensionally equivalent to our ordinary notion of inconsistency. The second is also a point made earlier in our discussion: by extending the notion of logical inconsistency as he does, the expressivist partly erases the distinction we make between moral error and logical error. Whereas we are perhaps willing to say that the person who disapproves of lying but approves of getting little brother to lie is making a moral mistake, we ordinarily aren’t willing to say that they are making a logical mistake. We don’t even ordinarily say that a person is logically mistaken if they both approve of lying and disapprove of it. We have, then, a second reason for thinking that further modification to the expressivist’s extended notion of inconsistency won’t make it extensionally equivalent to our ordinary notion of inconsistency. Third, because of its reliance on substantive moral claims, the expressivist’s notion of inconsistency for evaluative judgments is discontinuous with our notion of inconsistency for nonevaluative judgments. The expressivist will thus have to explain why we ought to work with two distinct notions of logical inconsistency; which notion do we employ when the set of sentences includes both evaluative and nonevaluative judgments? (It would seem that the expressivist must work with two distinct notions of inconsistency because, as van Roojen makes very clear (1996: 332, 334), when we are considering only nonevaluative judgments we want to be able to distinguish between logical inconsistency and pragmatic incoherence.)

Even if the expressivist can resolve all of these difficulties, there is still the matter of persuading us that his interpretation of the kinds of sample sentences that we saw in Chapter 3 (SS-II) is better than what the realist has to offer. With Hale’s modifications to the AC solution, we have a very complex system, or explanation. And the expressivist has had to give up trying to explain why we take inferences like ‘if P1 and P2, then C’ and ‘if (7) and (8), then (9)’ to be instances of modus ponens. What has the expressivist gained in return for this complexity? As Blackburn sees it, expressivism’s main virtue is being able to dispense with the realist’s additional metaphysical commitment. If moral values are real, the thought goes, they must be queer and unnatural (see, e.g., (Blackburn 1998: 84; 48 ff.)); realist moral properties, if they were to exist, would be incompatible with a scientific view of the world. But Blackburn doesn’t give any reasons for these claims that haven’t already been answered by the realist in Chapter 2.2[[28]](#footnote-28)8 Moreover, the expressivist rejects *both* of the realist’s core tenets. While there may be reason to be concerned about the realist’s mind-independence claim (the second core tenet), common moral beliefs and practices strongly suggest that we take there to be moral facts or truths: we think that we can be mistaken in our moral judgments, we think that some people are more morally perceptive than others, and so forth. So there is a whole range of phenomena that the expressivist will have trouble explaining because of his noncognitivism. This is perhaps why Blackburn tries to provide the expressivist with a substantial, but non-realist, notion of truth (1984: section 6.3).

There are, however, a host of problems with this substantial notion of truth (see, e.g., footnote 21 above). But we needn’t concern ourselves with these problems here. For even if the expressivist has a way to avoid all of them, he still won’t be out of the woods. That is because there is *a further dilemma* confronting him. Like Hale, I will refer to it as Wright’s dilemma, since it was Crispin Wright who first posed it (1988b). The dilemma is succintly described by Hale as follows:

*either* the q-realist attempt to earn for moral utterances the right to what Blackburn now calls ‘propositional surface’ will succeed, *or* it will not. If *not*, the project fails to provide a satisfactory anti-realist treatment of this region of discourse; but if it *does* succeed, it reinstates the notions of truth, assertion, etc., for statements in the region, with the effect that the q-realist winds up denying the very contrast—moral utterances as expressions of attitude rather than genuinely assertoric—in terms of which his projectivist construal gets off the ground. (Hale 1993b: 385; Hale’s emphasis)

In other words, if the quasi-realist expressivist succeeds with the project of providing a substantial irrealist notion of truth, the noncognitivist part of the project, it would seem, can be scrapped. For the expressivist’s real dispute with the realist, Wright argues, then regards the proper notion of truth to apply to our moral statements (Hale 1993a: 351), not whether those statements should be interpreted in a cognitive or noncognitive fashion.

Wright’s observation, I think, explains why it is that philosophers have often been puzzled by Blackburn’s quasi-realist project. The puzzlement arises from not being able to see how that project differs from the realist’s with respect to the cognitivism-noncognitivism issue. Blackburn writes:

The trouble is that the overall package [that the quasi-realist expressivist offers] is easily charged with incoherence. The charge is that it starts by depending on a contrast—that between attitudes and so forth on the one hand, and beliefs on the other. But it typically ends by assimilating them so closely that, it may be charged, the original contrast seems to evaporate. We understand the contrast at the beginning, so a critic might say, by thinking of beliefs, as opposed to attitudes, as capable of negation and disjunction, as sustaining truth, falsity, and . . . . But the forms of thought and language that embody these capabilities of belief are precisely the ones that the quasi-realist claims to earn for his attitudes. So in that case the original contrast collapses, and with it any coherent position. We cannot say ‘such-and-such commitments are not beliefs; they are other states that properly sustain the appearance of being beliefs’, if properly sustaining that appearance is all that there is to *being* a belief—all that there is to mark off beliefs from attitudes and the rest in the first place. (1993b: 366)

Blackburn responds to the charge by reaffirming that attitudes just are beliefs; attitudes, he says, are just those beliefs that are also attitudes (ibid., 367).2[[29]](#footnote-29)9 In this way, attitudes turn out to be a distinct kind of belief. Hale reports that Blackburn indicated to him in personal correspondence that “he is happy to see expressivism as a ‘throw-away-ladder’” (Hale 1993a: 360).3[[30]](#footnote-30)0 But I agree with Hale that if Blackburn tries to reinstate truth, he confronts Wright’s dilemma head-on.

So it would seem that the expressivist should settle for a deflationary, or minimalist, theory of truth, the only commitment being to Tarski’s T-schema. According to this view of truth, if I say “It is true that torturing small children is morally wrong” I am saying nothing more than “Torturing small children is morally wrong”.3[[31]](#footnote-31)1 The meanings of the two statements, this expressivist will say, are exactly the same. Hale argues that it is this *modest* form of quasi-realist expressivism (modest because the expressivist is not trying to reinstate a substantial notion of truth) that is viable as long as the expressivist can come up with the appropriate fixity operator for beliefs. But we have already seen that this expressivist is not able to explain why we think certain basic inferences involving evaluative judgments are instances of modus ponens; and this expressivist will have great trouble explaining a number of other important features of morality (e.g., how can the noncognitivist without a substantial notion of truth explain our belief that we can be mistaken in our moral judgments?). Furthermore, this expressivist still needs to come up with a fixity operator, still needs to resolve the problems associated with an extended notion of inconsistency, and will still need to address a number of other problems as well. In the next section I list several of these other problems.

This ends our discussion of the Frege-Geach objection. As yet, the expressivist is unable to adequately explain how we make use of evaluative and non-evaluative premises in our moral reasoning. The expressivist has to reject the Frege Point because evaluative predicates found in indirect contexts do not express attitudes of the speaker, or certainly not the same attitudes the speaker expresses when using these evaluative predicates in direct contexts. But rejecting the Frege Point leaves the expressivist in a bind when he tries to explain why we think certain basic inferences involving evaluative judgments are instances of modus ponens. It seems, then, that the very best the expressivist can do is try to explain why we take certain inferences to be valid and other inferences to involve a logical mistake. Yet we have also seen a number of reasons to think that the expressivist won’t be able to do this either. Even the expressivist’s most advanced proposal to date finds inferences to be valid that we wouldn’t ordinarily say are valid, and finds logical errors where we wouldn’t ordinarily find them.

The Frege-Geach objection illustrates the expressivist’s trouble making sense of the range of sentences involving evaluative predicates that we attribute to others, and the ways in which we employ these sentences (or their mental counterparts) in our moral reasoning. In this respect, the expressivist’s account of moral properties faces the same kind of problem as the sensation account of colour properties.

*Other Problems Facing Expressivist Accounts of Moral Properties*

1. The expressivist, I think, will not have the best explanation for our belief that human agents have free wills. We say things like

(J) Jones believes that Smith *chose* to steal the money and that is why he should be punished.

As Blackburn sees it (and I think Gibbard too), the main appeal of expressivism is its ability to reconcile ethics with the natural world (Blackburn 1998: Chapter 3). Ethics appears to present a problem in this respect because it is seen as a normative activity, something that takes place within “the space of reasons” rather than the domain of causality (ibid. 53). Thus, the expressivist aims to “place” ethics within a scientifically respectable view of the world. This aim arises out of a certain metaphysical assumption—the belief that a property is real, or “natural”, only if it can be detected by the human senses or investigated by the sciences of physics, chemistry, or biology (ibid., 48-49).3[[32]](#footnote-32)2 When Blackburn describes his position as a form of ethical naturalism, this is the kind of naturalism he has in mind.

However, this naturalism seems to entail a kind of determinism, one that denies that agents have free wills. We cannot reconcile ethics with the natural world unless we can “locate” the mental states that constitute our moral responses in a causal structure (ibid., 53). If, on the one hand, the expressivist locates free will within such a causal structure, it would seem that he is no longer talking about free will as we currently understand it. If, on the other hand, the expressivist argues for compatibilism between the notions of free will (as we currently understand it) and determinism, it would seem that he has given up on the goal of placing ethics within a scientifically respectable view of the world.

The determinism of expressivism is such that one needn’t worry, Blackburn tells us, about people learning that they actually *confer* moral value on things (1993a: 157, 176-77). Ordinarily we implicitly rely on our second-order (SO) beliefs to help us justify our first-order (FO) ethical views. For example, many of us hold the FO view that it is wrong to kick dogs just for fun. And many of us support that FO view with the belief that it is wrong because it causes the dogs unnecessary pain.3[[33]](#footnote-33)3 That belief in turn relies on a number of SO beliefs—such as the belief that it makes sense to justify our moral responses in terms of natural features; and it relies, I would argue, on the belief that we can be mistaken in our moral judgments. The expressivist, however, denies that there are any moral facts or truths; the reality, he says, is that we project moral value onto things. Thus, it isn’t really the case that it is wrong—not in any real sense—to kick dogs just for fun.

Of course many philosophers have worried, and rightly so, that were we to accept the truth of expressivism, many of our FO moral views and practices would change. For the expressivist, however, it is the determinism that prevents this from being the case. The moral “must” is categorical in the sense that we have no choice but to act as we do; we have no choice but to follow our sentiments: “The news comes in and the emotion comes out; nothing in human life could be or feel more categorical” (Blackburn 1993a: 177). Once this truth is recognized, “is there any substance left to the worry about failure of harmony of the theoretical and deliberative points of view?”, Blackburn rhetorically asks (ibid.). In this way Blackburn locates the mental states that constitute our moral responses in a causal structure, but at a huge cost. How does the expressivist explain our belief that agents can act freely?

2. Suppose the expressivist has a cogent response to the problem just posed; that is, suppose that he can consistently either deny any commitment to determinism or defend a form of compatibilism. There still remains the problem of explaining our reliance on our SO beliefs when justifying our FO ethical views, and the problem of explaining common SO beliefs about morality. Again, we commonly believe that we can be mistaken in our moral beliefs, we commonly believe that some people are more morally perceptive than others, and we commonly believe that we are capable of improving our ability to arrive at correct moral judgments. Expressivism doesn’t seem capable of adequately explaining the existence of these SO beliefs. It will thus also be true that the expressivist will have trouble explaining those of our FO beliefs and practices which “incorporate” such SO beliefs. (See, for example, issue 3. below.)

It seems right to say that we can distinguish between our ethical judgments or evaluations and the beliefs we rely on to justify those judgments or evaluations. In the area of art criticism, for example, we can ask three different levels of questions: (i) was the work of art good or bad?; (ii) what makes it good or bad? (that is, what are the criteria we are using to judge works of art?); and (iii) why these criteria and not some other? —which in turn forces the question: what constitutes a work of art? We can distinguish the same three levels in the area of moral criticism. Here I am assuming that a FO moral judgment is one which explicitly says whether some action or character trait is morally right or wrong, just or unjust, etc., and that a SO judgment is a certain kind of non-moral belief3[[34]](#footnote-34)4 upon which justification of our FO moral judgments depends in the way that level-(i) judgments in the area of art criticism depend on our level-(ii) and level-(iii) commitments. Or, as Brink puts it, we should think of SO issues as “issues about, rather than within, morality”; “First-order, or normative, issues, by contrast, are issues within morality about what sorts of things are morally important” (1989: 1).

Expressivists, and noncognitivists in general, want to maintain not only a sharp distinction between facts and values, but also a sharp distinction between FO (ethical) and SO (metaethical) views (Brink 1989: 3-5), (Singer 1973), (Blackburn 1993a: 173). We just saw in 1. why this makes good sense on the expressivist’s part. There is the worry that any moral agent who knows that moral value is merely projected onto the world by individuals will be inclined to change the way he talks about right and wrong, and should perhaps even change his behavior. If it really isn’t the case that kicking dogs just for fun is morally wrong (in some mind-independent sense), and one knows this, one would feel pressure to qualify their statement as follows: “I believe that kicking dogs . . . is morally wrong”; or: “I feel that it is wrong to kick dogs just for fun.” One will likely be less inclined to voice such feelings, too. For if there is nothing objectively wrong about kicking dogs gratuitously, there is less of a need to pressure those who are doing it to stop. Or, one might still want to pressure others to stop, but employ some other reason than the fact that it is morally wrong—since one doesn’t believe that it is morally wrong and those on the receiving end may also know enough to know that their actions aren’t really right or wrong. It would thus be far more honest and effective to say: “I don’t like it that you are doing that; please stop.”

Blackburn tries to avoid this difficulty by maintaining that any mental state which we employ in justifying our FO ethical views is itself first-order. He essentially denies that there are any SO issues:

[Question (put to Blackburn):] When philosophers ask whether moral truths are mind-dependent, the question appears to be asked at the meta-level . . .

[Answer:] The question *appears* to be asked at the meta-level, and perhaps is *intended* as a meta-level question. But there is no such meta-level question.

(1998: 311)

According to Blackburn, any statement we make within morality or about morality is a FO statement; “Talk of dependency [i.e., justification] is moral talk or nothing” (1993a: 173). It is this stance which allows him to sincerely say “Cruelty is not wrong because we disapprove of it, but because it causes pain and anguish” (1998: 74); and to say: “It is because of our responses that we *say* that cruelty is wrong, but it is not because of them that it is so . . . What makes cruelty abhorrent is not that it offends us, but all those hideous things that make it do so” (1993a: 172). The expressivist can say such things with sincerity because they are all being said at the FO level, and thus, they must all be interpreted expressivistically—which is to say, all they reflect is a further expression of the speaker’s attitude.

Blackburn also says: “This is not, of course, to deny that ‘external’ questions make sense—the projectivist plus quasi-realist package is an external philosophical theory about the nature of morality” (ibid.). Yet I don’t see how he can have it both ways. At some point we have to understand a sentence like “Cruelty is wrong because we disapprove of it” for what it is, and as it was intended to be understood; at the metaethical level, this is what the expressivist is telling us when he is saying that moral value is projected onto the world. If we are expressivists, then from the external, philosophical point of view, the view that cruelty is wrong because it causes anguish and pain has no greater authority than the view that cruelty is wrong because we disapprove of it.

So the dilemma for Blackburn is this: if we grant him his claim that “talk of dependency is moral talk or nothing,” and grant him his claim that he has an explanation for our SO beliefs about morality (they are just expressions of attitude and we can’t go any deeper than that), he is forced to either deny that we can make the kind of distinction between the FO and SO that we think we can make—a denial that is quite implausible, or to deny that our beliefs about the nature of morality (and any other matter that we would ordinarily say is SO) have any implications for our FO practices—a denial that is equally implausible. The other horn looks equally problematic: if the expressivist grants that our SO beliefs are indeed SO, he can avoid the implausible results of the first horn, but he now faces the problem of explaining those SO beliefs from an expressivist metaphysic, as well as the FO moral practices and beliefs which incorporate those SO beliefs. And I don’t see how the expressivist can give the best account of those beliefs.

3. One FO practice is to appeal to *natural features* in the world to justify certain of our FO ethical views. Again, many of us will say that it is wrong to kick dogs just for fun because it causes them unnecessary pain. In general, we assume that our moral judgments *supervene on* natural descriptions of the situations to which they apply. But the expressivist doesn’t seem to have a way to explain this supervenience claim.

He cannot say that this is simply what we do; he has to be able to make sense of the practice, especially in light of his denial that there are moral facts or truths. He has to say why it makes sense for us to think that agents *ought* to have certain responses to certain situations. In other words, the expressivist has to say why we think there are correct and incorrect moral responses, and why we think the correctness or incorrectness of these responses is partly determined by features of the situations to which we are responding. If we confer value on things, why must there be a fit between our moral responses and natural features in the world? Because the expressivist is a projectivist, he cannot say, as the colour sensationalist does, that certain features of situations *cause* us to have the moral responses that we do in fact have.3[[35]](#footnote-35)5 And the expressivist cannot say that certain features of situations justify our having the moral attitudes that we do in fact typically have. For if the features of situations are what justify our moral responses, it seems false to say that the primary function of moral discourse is to express attitudes of the speaker; it is far more accurate to say that we are responding to certain features of situations, and that we have the correct response if we have perceived all the relevant features and seen these features in the correct light; it is far more accurate to say that our moral judgments aim to be appropriate to the situations being judged.3[[36]](#footnote-36)6

But Blackburn explicitly says that this is not what the expressivist has in mind; the expressivist rejects “the view that when we moralize we respond to, and describe, an independent aspect of reality” (1993a: 157); the sentences that we use to voice our moral responses “express nothing to do with representing the world”; their “function is not to represent anything about the world” (ibid., 184-85).

Our belief in supervenience also seems to be incongruous with the idea that we are *projecting* attitudes on to things, and with the idea that the noncognitive response is primary and, as such, is the material out of which we construct the feature to which we think we are responding.3[[37]](#footnote-37)7 The expressivist will say that we think we are responding to the supervening property, but this is a property that doesn’t really exist in any objective way; it is what is projected onto the world. But the expressivist will then add to his story the idea that our moral responses are responses *to* natural features in the world,3[[38]](#footnote-38)8 features that some realists would consider to be the base properties upon which moral properties supervene. But if our moral utterances reflect a response to natural features in the world, why understand them in purely noncognitive terms? Why say that these utterances fail to represent anything about the world?

It is in this way that I see the expressivist’s noncognitivism to be incongruous with a commitment to supervenience. Supervenience requires us to say that the correct application of moral terms is a function of the subvenient properties.3[[39]](#footnote-39)9 Because determining, for example, whether an action is rude rather than cruel depends in part on the features of the action being judged, it is hard to dispute the claim that “ordinary evaluative thought presents itself as a matter of sensitivity to aspects of the world” (McDowell 1985: 110).

The expressivist seems fully conscious of the incompatibility between his noncognitivism and the supervenience thesis when he explains the existence of moral disagreements. Given the commitment to supervenience, one would think the explanation would run as follows: moral disagreements are possible because it is possible for people to express opposing attitudes toward the same action (or character trait, or some other appropriate “object”). Of course, a problem does arise for the expressivist when we ask, What is the disagreement really about? According to the understanding just given, the disagreement would seem to be about whether the object we are reacting to deserves the reaction we give it. Compare this explanation with the cognitivist’s: he will say that the argument is about whether the object we are judging deserves, or merits, the judgment we give it. In any case, this possible explanation for the expressivist is problematic because he claims that value is not a function of real properties in the world. As soon as we talk about the correctness of moral responses in terms of “fit” between response and features in the world, we make value a function of real features in the world. Thus, the expressivist offers a quite different explanation of moral disagreement. He explains it instead in terms of the disputants *wanting different things*, “both of which cannot happen” (Rachels 1999: 44).4[[40]](#footnote-40)0 While this response fits quite well with the expressivist’s commitment to noncognitivism, it looks to be very much incompatible with his commitment to supervenience and the idea that the correct application of moral terms is a function of the properties upon which they supervene. Supervenience suggests that the disputants are really disagreeing over what the best fit is between response and object (whether or not the former is understood in terms of desires, aims, goals, etc.).

One further observation adds to the need for the expressivist to defend the claim that his noncognitivism is compatible with the supervenience thesis. It is a problem for the expressivist that there fails to be a necessary connection between the *meaning* of a sentence (or word, or predicate) and the attitude expressed in certain *uses* of that sentence (or word, or predicate).4[[41]](#footnote-41)1 We saw in Chapter 2 that the expressivist will view the connection between moral claims and the natural claims4[[42]](#footnote-42)2 upon which they supervene as a *logically necessary* one. But the idea is not that we can logically deduce moral evaluative conclusions from premises that are all non-evaluative; it rather seems to be (or seems to have to be) that the very meaning of moral claims is to be “found in” the natural claims upon which they supervene. We are told, for example, that we are not competent users of moral terms unless we adhere to the supervenience constraint (Blackburn 1985b: 137). If the subvenient natural claims are not an outcome of an analysis of the meaning of the supervening moral claims, then it is not clear to me where the logically necessary connection comes from. This understanding of what gives us the logically necessary connection between moral claims and the natural claims upon which they supervene is not, by itself, a problem for the expressivist. But it becomes a problem for him when he identifies the meaning of moral utterances with their use—when the meaning is understood in terms of what the utterances can be used for. For it is the case that “there is more to meaning than use” (Searle 1962: 432). The gap between meaning and use further illustrates why there is a gap between the supervenience and noncognitivism theses. It is this gap that I am saying the expressivist needs to close if we are to be persuaded that the two theses are compatible.

4. Generally speaking, we understand our moral judgments to be evaluations. But in this case, it seems that there must be a standard, or criteria, against which we are evaluating, especially if we have any expectation of improving our moral judgments. My own view is that the criteria we rely on arise from our conception of morality (in the same way that what we think constitutes a work of art greatly influences the kind of criteria we use in judging whether particular works of art are good or bad). But, if there is a standard against which our moral evaluations are made, it makes sense to say that moral judgments ought to be interpreted descriptively. In other words, a full-fledged noncognitivism seems at odds with understanding our moral judgments as evaluations.

5. The expressivist will have trouble explaining attributions like:4[[43]](#footnote-43)3

(K) Jones believes that millions died in Russia as a result of Stalin’s inhumanity.

(L) Jones believes that Mother Teresa’s goodness won her a Nobel Prize.

It seems that certain general facts will not be identifiable or explicable unless we refer to moral properties: “. . . certain regularities—for instance, honesty’s engendering trust or justice’s commanding allegiance, or kindness’s encouraging friendship—are real regularities that are unidentifiable and inexplicable except by appeal to moral properties . . . Moral explanations allow us to isolate what it is about a person or an action or an institution that leads to its having the effects it does” (Sayre-McCord 1988c: 276). And because “these explanations rely on moral concepts that identify characteristics common to people, actions, and institutions that are uncapturable with finer-grained or differently structures categories” (ibid.), they are (a) extendable to new cases; and (b) cannot be supplanted, or easily supplanted, by explanations that don’t employ the moral concepts in question.

When we talk about Stalin’s inhumanity and Mother Teresa’s goodness, we seem to treat ‘inhumanity’ and ‘goodness’ as properties with a certain objective aspect to them. Sentences (K) and (L), for example, suggest that these properties have causal powers. How can the expressivist explain such indirect uses of these predicates? (‘Indirect’ in the sense that the predicates are again found in indirect contexts in (K) and (L).) The expressivist runs into trouble if he claims that we are really saying that Stalin was inhumane *because* he brought about the death of millions in Russia, for we are clearly not suggesting that Mother Teresa was good because she won the Nobel Prize. Stalin had certain dispositions, beliefs, and desires which led him to bring about the death of millions. So even if we say that someone is inhumane only in virtue of what they do, not in virtue of how they are disposed, we will still say that millions died in Russia as a result of Stalin’s morally flawed character. We will say: Stalin’s moral badness was the primary cause of millions dying in Russia during the 1930s; and: Mother Teresa’s goodness was the primary reason for her winning a Nobel Prize.

While the expressivist will concede that there are certain features that Stalin possessed which led to millions dying in Russia, he will insist that there is no such actual feature as moral badness. The only work that the predicate ‘moral badness’ is doing for us, the expressivist will say, is that of expressing an attitude toward those features possessed by Stalin which in fact brought about the death of millions. But while this interpretation maybe suffices for direct contexts, it leaves us in a bind again with the kind of indirect contexts we are dealing with here. For if ‘moral badness’ is functioning only to express an attitude, or primarily to express an attitude, it no longer makes sense to say what we say, or think what we think, in regards to Stalin. It doesn’t make sense to attribute causal powers to *our* attitudes toward Stalin, or certain features of Stalin. The expressivist cannot say that ‘moral badness’ refers to certain features of Stalin, not on a noncognitivist interpretation of the phrase. (Again, we see how noncognitivism’s incompatibility with the supervenience claim poses problems for the expressivist.)

David Brink builds on Sayre-McCord’s observations above regarding the explanatory potency of moral properties (Brink 1989: 193-197). I won’t give the details of Brink’s arguments here, but they do much to add to the case against the expressivist and the problems he will have explaining attributions like those of (K) and (L) above.

6. Finally, if the expressivist claims that his account is better than the realist’s where the explanatory criterion of reality is the standard, then he faces the following dilemma posed by Sayre-McCord:

Either there is a fact of the matter about which explanations are best, or there is not. If there is, then there are at least some evaluative facts (as to which explanations are better than others); if not, then the criterion will never find an application and so will support no argument against moral theory. (1988c: 277)

Sayre-McCord adds: “once it has been granted that some explanations are better than others, many obstacles to a defense of moral values disappear . . . [for] whatever ontological niche and epistemological credentials we find for explanatory values will presumably serve equally well for moral values” (ibid., 278). Sayre-McCord then proceeds to outline one type of argument that looks to be very promising in the defense of moral values.

It won’t do for the expressivist to say that we ought to interpret the explanatory criterion of reality expressivistically, although I suspect that this is exactly what he will want us to do. It won’t do because such an interpretation doesn’t capture what we intend to be saying when we say that one explanation is better than another. At some point the expressivist has to allow us to speak at the theoretical, or second-order, level. Otherwise he can never engage in a dispute with the realist regarding the metaphysical status of moral properties. Besides, we don’t think that people are merely expressing an attitude when they say that we can better explain many diseases in terms of bacteria and viruses than in terms of the afflicted person or creature being possessed by evil demons.

We can also combine this objection with those of 4. and 2. in the following way: if we understand morality to be an inquiry into how one should live, then moralizing seems to commit us to the existence of evaluative facts, for in virtue of engaging in this inquiry, we are assuming that some kinds of lives are better for us than other kinds of lives. But if the expressivist’s metaphysical claim is true, it doesn’t seem that we can make sense of this FO moral inquiry that commits us to the existence of these kinds of evaluative facts; the expressivist, it would seem, must resort to an error theory.

*Concluding Remarks*

The problems that the moral noncognitivist faces look to be insurmountable. Even if he finds workable solutions to all of them, there is little reason to think that he will have the *best explanation* of those common features of our moral beliefs and practices that need to be explained. The expressivist who is a quasi-realist not only has to explain the moral phenomena, but also our understanding of them. This means that the more subtle and elaborate his explanation of the former—that is, the more subtle and elaborate relative to our ordinary understanding—the more difficult it will be to explain why we have the understanding of those phenomena that we do.

It is therefore safe to conclude that we ought to accept something that is presupposed in the realist’s first core tenet, namely the idea that our moral utterances ought to be interpreted as having an important descriptive function. There are good reasons for understanding moral predicates as referring to, or describing, features in the world. The cost of not doing so is simply too high. Thus, if the antirealist is to succeed with his unmasking of realist moral beliefs and practices, he would do best to resort to a different strategy than that of noncognitivism.

1. 1 This realist appearance is due to features such as: the fact that we assert moral judgments in the same way that we make other kinds of assertions about what is so, the fact that moral predicates appear in embedded contexts, and the fact that our moral reasoning often presupposes that there are correct and incorrect answers to at least some moral questions. [↑](#footnote-ref-1)
2. 2 This is not to say that all noncognitivists hold that all moral utterances fail to have a descriptive function. But those noncognitivists who do say that moral utterances can have descriptive meaning hold that this descriptive meaning is always secondary in importance to the emotive meaning—i.e., the attitude expressed. At the end of this chapter I say a little about why I think this position is untenable. [↑](#footnote-ref-2)
3. 3 The error would be engaging in practices which assume the existence of real moral properties when no such properties exist, and the practices cannot be accounted for in antirealist terms. The expressivist can say that no such error exists if he can account for the apparently realist practices. [↑](#footnote-ref-3)
4. 4 In this chapter I am concerned only with the expressivist who is a quasi-realist. In Chapter 5 I address the error theorist, although one who is a moral cognitivist. I am not sure how to address the expressivist who is an error theorist because I am not sure which common features of our moral beliefs and practices he will explain as error-free and which he will say are in error. If he says we are not in error when employing evaluative predicates in embedded contexts, then he faces the Frege-Geach objection discussed below in addition to the kind of problem discussed in Chapter 5—that of disinguishing between moral responses, and distinguishing moral responses from other kinds of noncognitive responses. [↑](#footnote-ref-4)
5. 5 That is, enough for the purposes of arriving at a conclusion about the prospects of expressivism, and enough in terms of the effort required for a proper treatment of that objection. [↑](#footnote-ref-5)
6. 6 More precisely, this understanding seems to commit us to the cognitivist aspect of the realist’s first core tenet. [↑](#footnote-ref-6)
7. 7 Again, this is because one can sincerely affirm P1 without holding the attitudes that are said to be expressed when the components of P1 are directly asserted. [↑](#footnote-ref-7)
8. 8 Geach first points to this problem in a footnote in “Imperative and Deontic Logic,” *Analysis* 18 (1958): 49-56. He addresses the problem in detail in “Assertion,” *Philosophical Review* 74 (1965): 449-465. John Searle also discusses the problem in (Searle 1962). Searle’s paper emphasizes the need to distinguish the meaning of sentences from their use. On this issue, see also pp. 13-14 of (MacIntyre 1981). [↑](#footnote-ref-8)
9. 9 Obviously there will be problems with any inference that we ordinarily make and see as valid when the inference is from sentences containing an evaluative predicate which appears in both direct and indirect (embedded) contexts. [↑](#footnote-ref-9)
10. 10 See (Schueler 1988), (Hale 1986), (Brighouse 1990), and (Zwangwill 1992). [↑](#footnote-ref-10)
11. 11 I am stealing this very succinct phrase from van Roojen (1996: 311). [↑](#footnote-ref-11)
12. 12 At the beginning of “Attitudes and Contents” Blackburn argues that modus ponens needn’t work in the way that van Roojen mentions. But in making that argument, Blackburn still assumes that a modus ponens argument will have the form ‘if *P* and *P* → *Q*, then *Q*’. But neither the STW interpretation nor the MR interpretation of our example argument give the argument that form. Thus, even according to Blackburn’s own understanding of modus ponens, the STW and MR solutions fail to preserve it (regarding the claim about the STW solution, see (Blackburn 1993a: 188-89)). [↑](#footnote-ref-12)
13. 13 It won’t do for a number of reasons. First, our understanding of the inference in question doesn’t seem to have come about due to explicitly held realist assumptions; we cannot conclude that a person consciously subscribes to a realist metaphysic if they understand the inference in the ordinary sense. Second, it would be a mistake to try to explain our first-order (FO) moral judgments completely separately from our second-order (SO) moral judgments when we rely so heavily on the latter to guide us in making the former. (Here I am assuming that a FO moral judgment is one which explicitly says whether some action or character trait is morally right or wrong, just or unjust, etc., and a SO judgment would be a certain kind of non-moral belief that can be used to try to justify our FO moral judgments. SO judgments will state beliefs about the standards we should rely on when making FO moral judgments, and they will state beliefs about the nature of morality. Norman Dahl adds: “SO beliefs would seem to be beliefs about FO beliefs. They would include some understanding of what counts as a justification, or what the relation of justification amounts to.”) Of course, if the quasi-realist does go this route, he can take a pass on explaining all sorts of moral beliefs: e.g., the belief that some of our moral judgments can clearly be correct or incorrect; the belief that we can improve in making moral judgments; a belief of the sort that it is wrong to kick dogs just for fun because it causes them unnecessary pain; the belief that a good part of morality has to do with benefits and harms to human beings; and so forth. But then it is hard to put much stock in the quasi-realist’s project. If our SO practice is potentially in great error, then it must be that our FO practice is too, seeing how the two levels of moral thought are so deeply intertwined. Finally, I think the *expressivist* quasi-realist runs into even greater problems if he tries to make such a reply. The expressivist will want to say that we should treat our SO statements expressivistically (Blackburn 1993a: 153); it would simply be implausible to say that the utterance “It is wrong to kick dogs just for fun” is an expression of an attitude but the utterance “It is wrong to kick dogs just for fun because it causes them unnecessary pain” is not. But then it follows that it would be very odd to say that in the one case our practice is not in error, but in the other case it is. (Under Blackburn’s expressivism, anything we rely on to justify our FO moral claims is also part of FO morality. I discuss some of the problems with this view much later in the chapter.) [↑](#footnote-ref-13)
14. 14 I think that Blackburn wants us to assume that denying the conclusion, C, amounts to not having an attitude of disapproval toward getting little brother to tell lies. So if one tolerates or approves of getting little brother to tell lies, one is “denying” the conclusion. Since I want to avoid employing the scare-quotes every time, the reader should keep in mind that when the sentences are interpreted expressivistically, the story about what we are doing and why we think the inference is justified is importantly different from our ordinary understanding of what is going on. [↑](#footnote-ref-14)
15. 15 Ordinarily, we say that a set of sentences is inconsistent if we can derive a contradiction from them, that is, if we can derive a sentence of the form *‘P* and not-*P*’; the problem with a sentence of this form is that it is false in all logically possible worlds. [↑](#footnote-ref-15)
16. 16 Even so, the expressivist’s account cannot be too different from our own understanding of what we are doing when making the types of inferences in question. Otherwise the expressivist won’t be able to plausibly explain why we see things as we do. [↑](#footnote-ref-16)
17. 17 Of course, the expectation is no longer that Blackburn show how the inference to C is valid per our ordinary understanding of logical validity, nor that Blackburn show how the inference is an instance of modus ponens as we ordinarily understand that inference rule. In “Attitudes and Contents” Blackburn explicitly says that he believes that the STW solution preserves modus ponens and validity. [↑](#footnote-ref-17)
18. 18 This is another way in which Blackburn glosses the semi-colon that is employed in Eex. [↑](#footnote-ref-18)
19. 19 The latter is something he has to do because we rely on this understanding to help us justify our first-order moral judgments, and the expressivist sees anything we use to justify our first-order judgments as being part of first-order moral practice (see, e.g., (Blackburn 1993a: 153)). [↑](#footnote-ref-19)
20. 20 Here one might think of numerous American politicians who give the highest praise to those who fight America’s wars, but who easily tolerate the fact that their own children are not in the military or that they themselves never served in the military. [↑](#footnote-ref-20)
21. 21 Later on in Chapter 6 of *Spreading the Word* Blackburn resorts to the notion of a “unique best possible sensibility” (198), but one in which “truth belongs to the trunk” (199). But he doesn’t say why the expressivist is justified in resorting to such a notion. (If he says that ‘best’ is simply functioning expressively here, then from a purely descriptive standpoint he is talking only about a unique sensibility, and this notion won’t do the work that he clearly wants it to do.) [↑](#footnote-ref-21)
22. 22 J. Hintakka, “Deontic Logic and Its Philosophical Morals’, *Models for Modalities* (Dordrecht: Reidel, 1969). [↑](#footnote-ref-22)
23. 23 As I see it, this move greatly helps the plausibility of Blackburn’s talk about inconsistent attitudes, since it is a tautology that two goals that conflict cannot both be realized. However, I think that we need to be careful before accepting this understanding of all moral attitudes. Is Blackburn now abandoning his notion of a fractured sensibility? Are the only dispositions that are morally important those which are goal-oriented and forward-looking? If what is centrally important about moral attitudes are the goals that they express, wouldn’t it be better to talk directly in terms of aims and goals? In what way does it remain important that moral responses are affective responses? Do all moral attitudes have a direct practical counterpart? That is, can all moral attitudes be realized in just one way? Are all aims and goals moral ones? If so, why? If not, why not? Since it seems unlikely that all moral attitudes will be realized in just one way (one would think that the “realization” of a moral attitude will depend on more than the attitude itself; one would think that the practical import of the attitude will also be a function of the agent’s belief states), it will be very difficult, I think, to determine whether two particular attitudes are inconsistent. As we will see below, an agent’s beliefs will also play a role in determining whether two attitudes are inconsistent. [↑](#footnote-ref-23)
24. 24 In AC, Blackburn tells us that his STW and AC solutions to the problem of indirect contexts both preserve modus ponens (1993a: 182-83, 186, 188-89) and that both also preserve the conventional notion of validity (186). This is not to say that the inference from P1 and P2 to C, under either expressivistic interpretation, is an instance of ordinary modus ponens, or that the inference is valid in the conventional sense. It is to say that the expressivist has earned “our right to propositional forms—including the use of a truth predicate” (186).

    Truth, for Blackburn, will be understood in a non-representational sense (184). While Blackburn sometimes favors a minimalist, disquotational notion of truth (1998: 78-79), he also tries to construct a substantial notion of truth (STW, Ch. 6, 198-199). Regarding the latter, Blackburn suggests that we understand the correctness or incorrectness of a particular moral response—say, to circumstances C—in terms of whether it is a response that someone with a “unique best possible sensibility” would have to C, or could have to C (since truth only “belongs to the trunk” (199)). With regard to moral statements themselves, Hale points out that the notion of moral truth which Blackburn appears to rely on is the following: “a moral statement is true if the attitude it expresses belongs to the limiting set of attitudes that results from taking all possible opportunities for improvement, etc.” (1993a: 360).

    A number of problems arise if Blackburn’s notion of validity is supposed to be truth-preserving in the latter sense. For example, such an understanding of validity doesn’t square with our everyday understanding of what is involved in a modus ponens inference, for that everyday understanding has nothing to do with a unique best possible sensibility. How does the expressivist explain this? Until such an explanation is provided, we have to worry whether Blackburn’s STW and AC solutions really preserve modus ponens. Also, how is this substantial notion of truth connected to the notion of ‘joint satisfiability’ that helps us judge the validity of arguments in the AC solution (1993a: 189; Hale also points to this problem in his 1993a article (p. 350))? [↑](#footnote-ref-24)
25. 25 This is a close paraphrase of Blackburn’s definition at (1993a: 194). I apologize for the use of both ‘~’ and ‘¬’ to denote negation. Hale favors the latter, Blackburn the former. In the present discussion, the reader should see the symbols as interchangeable. [↑](#footnote-ref-25)
26. 26 Again, I think we should be wary of this interpretation. ‘Approving of giving to famine relief’ is quite different from ‘not tolerating not giving to famine relief’. [↑](#footnote-ref-26)
27. 27 Here Blackburn refers to the functionalism that he discusses on pp. 56-57 of *Ruling Passions*. We are to understand his brand of expressivism as a form of “non-descriptive functionalism” or “practical functionalism” (1998: 77). [↑](#footnote-ref-27)
28. 28 See, for example, section 6.1 of (Blackburn 1984). [↑](#footnote-ref-28)
29. 29 Blackburn finds himself in this position because he has a functionalist view of mental states, and his quasi-realist claim is that moral attitudes properly sustain the appearance of being beliefs; yet, all there is to being a belief, from a functionalist point of view, is properly sustaining the appearance of being a belief, in functionalist terms. [↑](#footnote-ref-29)
30. 30 It is also worth noting what Blackburn says regarding his substantial notion of truth on p. 313 of *Ruling Passions*. [↑](#footnote-ref-30)
31. 31 See, for example, (Blackburn 1998: 79). [↑](#footnote-ref-31)
32. 32 Blackburn suggests in a number of places that ethical practices and states of mind are explicable in evolutionary terms ((1993a: 168-69, 205); although see (1998: 121)). This is something that another major proponent of expressivism, Allan Gibbard, also proposes (1990). The suggestion, in part, is that coordination among human beings increases our chances of survival, and thus from an evolutionary standpoint our moral behavior makes sense. I think it is very clear, however, that morality, as we presently understand it, is quite at odds with the evolutionary fitness of the human species. For example, we presently take there to be moral injunctions against killing individuals who are temperamentally unfit to function well in society, against polygamy and eugenics, against killing the elderly so as not to burden society or put too great a demand on the earth’s resources, against taking away the reproductive rights of those who are mentally or physically infirm, etc., etc. [↑](#footnote-ref-32)
33. 33 I am inclined to say that the belief expressed by the sentence “Kicking dogs just for fun is morally wrong because it causes them unnecessary pain” is a SO belief, although I suspect most philosophers would disagree. As Brink observes, there is room for disagreement about where we should draw the line between the first-order and the second-order (Brink 1989: 1). Although the sentence just referred to is not the same as “Unnecessary pain is a morally bad thing”, one might say that the mental states that these sentences are meant to reflect are so closely connected that they cannot be pulled apart. This would be the reason for saying that the first belief is FO. [↑](#footnote-ref-33)
34. 34 The adjective ‘non-moral’ isn’t especially helpful if we want to say that the belief expressed by the sentence, “Kicking dogs just for fun is morally wrong because it causes them unnecessary pain”, is a SO belief. In my view, we cannot draw a sharp distinction between moral and non-moral beliefs. [↑](#footnote-ref-34)
35. 35 See (Blackburn 1993a: 204-05). Here Blackburn responds to a similar charge brought by Nicholas Sturgeon regarding the projectivist’s inability to explain our explaining certain events by referring to moral properties which we take to be real. I address this related charge below. [↑](#footnote-ref-35)
36. 36 Norman Dahl points out that the expressivist is likely to respond as follows: “the expressivist can say that given that the purpose of moral talk and thought is to provide guidance that people can act on and provide it on the basis of principles that we would accept or approve that everyone act on, one won’t be able to provide such guidance except in virtue of principles that call attention to certain non-moral features of the world that in principle everyone could recognize. That is, it is an understanding of the purpose of moral speech and language that allows one to see why moral matters supervene on non-moral matters.” To my mind, however, this response is very much incomplete. The expressivist has to say much more about the purpose of moral talk and thought before it is clear why moral matters should be understood as supervening on non-moral matters. The expressivist is saying that the correctness of a moral “judgment” or response needn’t be understood in terms of the fit between that response and features of the object to which one is responding; rather, he seems to suggest that we can understand the appropriateness of a moral response in terms of how well it accords with the purposes for which we evaluate things. (Blackburn actually offers a substantial notion of truth in terms of a unique best possible sensibility. But a unique best possible sensibility, I take it, is one which best fits the purposes for which we evaluate things.) But what is the connection between the purposes for which we evaluate things and features of the objects to which we are responding? Until the expressivist answers this question, it is not at all clear why moral matters should be understood as supervening on non-moral matters. In addition, when the expressivist tells us the story of the connection between the purposes for which we evaluate things and features of the objects to which we are responding, he has to avoid resorting to cognitivism. Finally, as Norman Dahl observes, the expressivist’s account of (the purpose of) moral talk and thought must fit our actual moral talk and thought. We have seen in our discussion of the Frege-Geach objection, however, that the best current account that the expressivist offers doesn’t fit. So we have further reason for saying that it is not at all clear how the expressivist’s response alluding to the purpose of moral talk and thought helps to explain why we are justified in thinking that moral matters supervene on non-moral matters. [↑](#footnote-ref-36)
37. 37 See the passage quoted from Warren Quinn in the section at the beginning of Chapter 2 entitled ‘*The moral realist and his opponents, briefly characterized*’. [↑](#footnote-ref-37)
38. 38 Blackburn talks about “inputs” and “outputs” in section 1.2 of *Ruling Passions*. [↑](#footnote-ref-38)
39. 39 This is accepted by both realist and antirealist. See, e.g., (Sayre-McCord 1988c: 275, n. 43), (Blackburn 1993b: 378), and (Blackburn 1993a: 159). [↑](#footnote-ref-39)
40. 40 See also (Blackburn 1998: 69). The approach here, again, is to understand moral attitudes in terms of aims or goals. [↑](#footnote-ref-40)
41. 41 See, for example, (Searle 1962). Also, Alasdair MacIntyre observes that

    The emotive theory . . . purports to be a theory about the meaning of sentences; but the expression of feeling or attitude is characteristically a function not of the meaning of sentences, but of their use on particular occasions. The angry schoolmaster, to use one of Gilbert Ryle’s examples, may vent his feelings by shouting at the small boy who has just made an arithmetical mistake, “Seven times seven equals forty-nine!” But the use of this sentence to express feelings or attitudes has nothing whatsoever to do with its meaning. This suggests that we should . . . consider whether [the emotive theory] ought not to have been proposed as a theory about the *use*—understood as purpose or function—of members of a certain class of expressions rather than about their *meaning*—understood as including all that Frege intended by ‘sense’ and ‘reference’. (MacIntyre 1984: 13) [↑](#footnote-ref-41)
42. 42 Natural claims being claims that describe, in natural property terms, the situations to which the moral claims are being applied. [↑](#footnote-ref-42)
43. 43 Regarding both (K) and (L), see (Sayre-McCord 1988c: 275). [↑](#footnote-ref-43)