Moral Expressivism and Sentential Negation

Neil Sinclair (neil.sinclair@nottingham.ac.uk)

NOTE: This is an unrevised draft of the paper that appeared in *Philosophical Studies* 152(3) 2011, pp. 385-411. The original publication is available at www.springerlink.com and http://www.springerlink.com/content/j462420335564150/fulltext.pdf

Abstract

This paper advances three necessary conditions on a successful account of sentential negation. First, the ability to explain the constancy of sentential meaning across negated and unnegated contexts (the Fregean Condition). Second, the ability to explain why sentences and their negations are inconsistent, and inconsistent in virtue of the meaning of negation (the Semantic Condition). Third, the ability of the account to generalize regardless of the topic of the negated sentence (the Generality Condition). The paper discusses three accounts of negation available to moral expressivists. The first – the 'dominant commitment account' – fails to meet the Fregean Condition. The two remaining accounts – one suggested by commitment semantics and the other by recent analyses of the 'expression' relation – satisfy all three conditions. Mark Schroeder's argument that the dominant commitment account is the only option available to expressivists is considered and rejected.

Can expressivists give an account of the meaning of sentential negation? The initial problem is that the most natural way to understand negation is in terms of truth-conditions: If the meaning of an indicative sentence is given by the conditions under which it is true, then the meaning of its negation can be discerned by understanding that a negated sentence is true just in case the unnegated sentence is false. For expressivists the underlying semantics for moral sentences is not truth-conditional; they must therefore depart from the natural understanding of negation, at least in moral contexts. The problem is not so much the unnaturalness of this departure (for the fact that a second-order theory, even a semantic theory, comes naturally to us should bear little argumentative weight) but the very possibility of an expressivist position that captures all the semantic properties of negation.

In this paper I discuss three ways for expressivists to address their problem with sentential negation (hereafter just 'negation'). The first of these accounts is suggested by Schroeder (2008a, 2008b), who argues not only that this account solves the problem, but that it is the only way for expressivists to do so. I argue that Schroeder is wrong on both counts. I offer two alternative expressivist accounts of negation, both of which avoid a fatal flaw apparent in Schroeder's account, and are otherwise plausible.[1] In the penultimate section, I diagnose how Schroeder was led astray. The conclusion is that the negation problem fails to threaten the viability of the expressivism, although expressivists are well-advised to take a different path to the one Schroeder suggests.

1. Expressivism and Embedding

Expressivists' problem with negation is an instance of the more general problem of sentential embedding. Expressivism starts as an account of the meaning of simple moral sentences, that is, sentences of the form 'x is M' where M is a moral predicate and x is a subject non-morally described. According to expressivists these sentences express non-cognitive attitudes directed at x (and perhaps also at attitudes towards x). The attitude expressed, and the co-operative action-guiding nature of this expression, gives the meaning of the sentence. Expressivists typically give a structurally identical semantic story for simple non-moral sentences, only these are taken to express beliefs rather than attitudes.[2]

The embedding problem arises because natural language is more complex than this basic theory allows. All simple sentences can be embedded in contexts where they form just part of a larger sentence. For example, 'x is M' can occur as a sub clause in the sentences: 'It is not the case that x is M', 'It is false that x is M', 'John believes that x is M' and 'x is M because y is N'. The embedding problem is the problem of providing an account of the sentences formed by such embedding contexts – call them complex sentences – and hence an account of the meaning of sentences as they appear as sub clauses in such contexts.[3] The most discussed type of embedding problem concerns indirect contexts where the embedded sentence is no longer asserted. The negation problem falls into this subclass.

What are the conditions on a solution to the negation problem? The following three conditions can be abstracted from the recent literature.[4]

First, an account of negation must explain how the meaning of sentences, both simple and complex, remains constant across negated and unnegated contexts. 'x is M' means the same when it occurs in the sentence 'x is M' as when it occurs inside the sentential negation operator in '¬(x is M)'. A simple proof of this is that the two sentences are inconsistent. I shall call this the Fregean Condition, following Geach (1965). It follows from the Fregean Condition that the meaning of a negated sentence is some function of the meaning of the sentence it negates.

Second, an account of negation must explain why systematic semantic relations hold between sentences and their negations, and hold in virtue of the meaning of the negation. Here the relevant semantic relation is the one appealed to above: the inconsistency of 'x is M' with '¬(x is M)'. It is in virtue of the meaning of the operator '¬' that these sentences are inconsistent.[5] Hence an account of negation must explain this inconsistency, and explain it by reference to the semantic contribution of negation. Call this the Semantic Condition.

Third, an account of negation must be generalizable regardless of the topic of the sentence embedded. The sentence 'x is Φ' can be sensibly negated regardless of whether 'Φ' is a moral or non-moral predicate. Further, complex sentences such as conditionals, disjunctions and propositional attitude ascriptions can also be negated. Across all these contexts, the distinctive semantic contribution of negation remains constant (at least, common understanding sees no difference between the functioning of negation in these contexts). An account of negation needs to preserve this univocality. Call this the Generality Condition.

Satisfying each of these conditions is arguably necessary for a successful account of negation. The argument is that any account that didn't satisfy one of these conditions would be committed to denying obvious semantic facts, such as the fact that sentences preserve their meaning when embedded, the fact that sentences and their negations are inconsistent or the fact that negation is univocal. I will not argue that joint satisfaction of these conditions is sufficient for a successful account of negation. For all that is said here, it remains a possibility that there are further conditions on such an account. But no claim of sufficiency is necessary for my argument. Of the three accounts of negation offered I argue that the first fails to meet the Fregean Condition and therefore should be rejected. The second and third meet all three conditions.

1. The Unwin/Hale Problem

Before coming to these accounts, it is worthwhile to discuss an argument developed separately by Unwin (1999) and Hale (2002). Though this argument can be presented as an argument against the very possibility of an expressivist account of negation, it is better understood as providing an indication of the sort of mistakes that expressivists needs to avoid.

Consider an agent who accepts that x is M. We might consider three contrasting characters. First, an agent who accepts that not-x is M. Second, an agent who accepts that: it is not the case that x is M. Third, an agent who simply has no opinion as to whether x is M or not-M. We can represent these positions thus:

(1) 'x is M'
(2) '¬x is M'
(3) '¬(x is M)'
(4) 'Harrumph'

There are a couple of points worth noting here.

First, in (2), '¬' functions in a perfectly understandable way as a subject-operator rather than a sentential-operator. This is sometimes called internal negation to contrast it with the external negation evident in (3) and which is our primary

concern. Internal negation is easily understood when the subject is a gerund such as 'murder', where '¬x' is 'not murdering'. It can also, somewhat less naturally, be understood where x is an object like 'the girl': here '¬x' would roughly translate as 'everything but the girl'. In what follows, I accept the common assumption that '¬x' makes sense given either type of substitution for x.

Second, the position expressed in (4) is the position of someone who has no opinion as to the M-ness of x, although she may have opinions about other issues. (Perhaps this is better represented as 'Harrumph$_{\text{M-ness of x}}$', but I avoid this unnecessary complexity). This position covers two distinct possibilities, which we may stipulate as Indifference and Agnosticism. The Indifferent agent simply has no opinion regarding the M-ness of x, either because she hasn't yet considered the matter or because she is yet to confront what she considers sufficient reason for assent or dissent. By contrast, the Agnostic agent holds that there can in fact be no sufficient reason for assent or dissent (or, more strongly, that there could not possibly be such a reason). Both the Indifferent agent and the Agnostic avoid the issue of the M-ness of x, but they do so in different ways: the Agnostic has ruled herself out ever assenting (or dissenting) to the claim that x is M in a way that the Indifferent agent has not. (The Agnostic is a particular concern of Hale (2002)).

The Unwin/Hale problem is as follows. According to expressivists (1) expresses an attitude directed at x (and perhaps also at attitudes directed at x). For our purposes it doesn't matter whether this attitude is positive or negative, so it can be formalized as 'α!x'. The problem is giving an expressivist account of (3). Following the expressivists' central claim, we might imagine that the meaning of (3) is given by the attitude it expresses. Further, since we are seeking an account of negation, we might imagine that the best way to give an account of this attitude is to insert a negation into the description of the attitude expressed by (1), namely 'α!x'. The problem is that it seems there are only two places this negation can go and both result in the conflation of (3) with one of the other of the positions contrasted with (1).

First, suppose the negation appears inside the attitude-operator, so that (3) expresses the attitude 'α!¬x'. The problem is that this view it conflates (3) with (2), since if (1) expresses α!x then simple substitution guarantees that (2) expresses α!¬x.

Second, suppose the negation appears outside the attitude-operator, so that (3) expresses the attitude '¬α!x'. Now it is initially unclear what the symbol '¬' might signify when it occurs in front of an attitude-operator, but two possibilities suggest

themselves. First, perhaps '¬α!x' is simply an absence of the attitude 'α!x'. But this cannot be what is expressed in (3), for it conflates the position expressed there with Indifference. Second, perhaps '¬α!x' indicates an ongoing commitment not to have the attitude 'α!x'. But this also cannot be what is expressed by (3), since it conflates the position expressed there with Agnosticism. If these are the only possible interpretations of '¬' is it occurs outside an attitude-operator, this version of expressivism is committed to conflating the position expressed by (3) with one or other of the positions expressed by (4). Either way, therefore, expressivism conflates positions that should not be conflated.

This problem is not fatal to expressivism.[6] The proper conclusion is only that if the meaning of (3) is given by the attitude it expresses, this attitude cannot be the same as any of the attitudes expressed by (2) or (4). It is true that an account of this attitude cannot be easily generated by application of the negation symbol in the ways Unwin and Hale consider. But it doesn't follow that no such account can be given. Two such accounts are given below, in §§2-3. A third option is to deny the antecedent of the troublesome conditional, and this is discussed in §4.

2. The Dominant Commitment Account

The Unwin/Hale problem is that in moving from the attitude expressed from an unnegated sentence to the attitude expressed by the negated sentence, there are simply not enough places for negation. The obvious solution is to discern some complexity in the attitude expressed by the unnegated sentence that generates a further place for negation.[7] What sort of complexity might this be? One suggestion is to borrow from expressivist accounts of the distinctive nature of the moral attitude. On a well-known view (which has been parenthetically trailed) moral attitudes are distinctive insofar as they are ascended, that is, directed not only at objects but at attitudes towards those objects.[8] Suppose we focus on the latter and say that 'x is M' expresses an attitude directed at attitudes towards x. We can formalize this as 'α!β!x'. The extra attitude-operator generates an extra place for negation, opening up the possibility of a satisfactory account of the negation-attitude. Schematically, we can distinguish the attitudes expressed in our four positions thus:

(1') 'x is M'	expresses	α!β!x
(2') '¬x is M'	expresses	α!β!¬x
(3') '¬(x is M)'	expresses	α!¬β!x
(4') 'Harrumph'	expresses	¬α!β!x

As an example, take 'x is M' to be 'Murder is wrong'. Schroeder (2008a, 589) suggests that 'α!' might be the attitude of 'being for', which is a 'very general positive attitude' and that 'β!' might be the attitude of 'blaming for', understood in the usual way. Hence 'Murder is wrong' expresses *being for: blaming for murder*, whereas 'It is not the case that murder is wrong' expresses *being for: not blaming for murdering*. The latter attitude is distinct from the attitude expressed by 'Not murdering is wrong', which is *being for: blaming for not murdering*. It is also distinct from the attitude of those who are Indifferent or Agnostic about the wrongness of murder: the former is a mere absence of any relevant attitude of being for, the latter is a considered commitment to the absence of such an attitude. Hence the Unwin/Hale problem is solved.

This solution generalizes to all simple moral sentences. It doesn't rely on these particular examples of 'α!' and 'β!' – we may suppose that different moral predicates

are associated with different attitude pairs. All the view requires is that the attitudes expressed by simple moral sentences exhibit enough complexity to warrant three negations: two inside to the dominant attitude-operator and one outside. Furthermore, this view mirrors a simple account of the semantic role of negation when the unembedded sentences are taken to express beliefs rather than attitudes. Since the objects of beliefs are structured propositions rather than objects, beliefs have two places for negation within the scope of the belief-operator: one that modifies the subject of the proposition and one that modifies the whole proposition. Thus the Unwin/Hale problem is not a problem for sentences expressive of belief. The dominant commitment account shows, in just the same way, why it needn't be a problem for sentences expressive of attitude (Schroeder, 2008a, 590).

Call this suggestion the dominant commitment account, since it supposes that both negated and unnegated simple sentences express the same commitment type – the dominant commitment – but directed at different objects. Here 'commitment' is intended in the usual sense that includes both beliefs and attitudes and may include other types of mental state, such as personal projects, patterns of emotional reaction and so on. In the case of simple moral sentences, the dominant commitment type is an attitude ('α!') and its objects are attitudes directed at objects or actions ('β!x').

a. Semantic Condition

How does the dominant commitment account fare with respect to the three conditions on an account of negation?

Take first the Semantic Condition. Does the account explain the inconsistency of 'x is M' and '¬(x is M)' in the right sort of way? According to Schroeder (2008a, 593) the explanation has two components. The two sentences are inconsistent because they express (i) tokens of an inconsistency-transmitting commitment type, namely 'α!', that are (ii) directed at inconsistent contents, namely the contents 'β!x' and '¬β!x'. Both points require some explication.

First, a commitment type is inconsistency-transmitting just in case having two tokens of that commitment-type directed at inconsistent contents is inconsistent. The paradigm example of an inconsistency-transmitting attitude-type is belief: believing that x is F and believing ¬(x is F) is inconsistent. Schroeder (2008a, 581) labels this type of psychological inconsistency 'A-type'. According to the dominant commitment

account, the sentential inconsistency of negated and unnegated sentences is the result of the A-type inconsistency of the states of mind they express. By extension from Schroeder's terminology, we can call this form of explanation the 'A-type model'.

Second, two contents are inconsistent when they cannot be simultaneously realized. This definition applies both to propositions (the contents of beliefs) and objects, actions and attitudes such as 'blaming for murder' (the contents of attitudes). In the former case, the propositions 'Grass is green' and 'It is not the case that: grass is green' cannot be simultaneously realized because they cannot both obtain. In the latter case the attitudes 'blaming for murder' and 'not blaming for murder' cannot be simultaneously realized because the same agent cannot do both. In both cases, the contents are inconsistent in so far as they cannot be simultaneously realized.[9]

To see how the A-type model works in practice, take an example. Suppose that 'Murder is wrong' expresses *being for: blaming for murder* and that 'It is not the case that: murder is wrong' expresses *being for*: *not blaming for murder*. On the dominant commitment account, the sentences are inconsistent because the commitments they express are psychologically inconsistent and this in turn is because (i) 'being for' is an inconsistency transmitting commitment-type and (ii) 'blaming for murder' and 'not blaming for murder' are inconsistent contents.

Is this a good explanation of the sentential inconsistency? There is one gap. The expressivist will need to explain why certain commitment types, such as *being for*, are inconsistency-transmitting. Schroeder sounds a note of optimism for expressivists here, arguing that

> It is intelligible for expressivists to hope that whatever explains the inconsistency-transmitting character of belief…will also explain why [moral attitudes] are inconsistency-transmitting (Schroeder, 2008a, 577).

Whether this promissory note can be cashed will be considered latter (§3c). For now, at least, there is no reason to suppose that the account of psychological inconsistency cannot be transferred to the dominant commitment type involved in the moral case.

There is, however, a potential objection to the A-type model. The objection is the old one that any account of complex logical sentences, such as negations, that supposes those sentences' logical form to be one where an attitude-operator is dominant cannot construe failures of inference involving those sentences as involving

logical, as opposed to mere attitudinal, failings.[10] Since the dominant commitment account just is the view that, when simple moral sentences are negated, the proper logical form has a dominant attitude operator, it seems to fall foul of this objection: having the attitudes 'α!β!x' and 'α!¬β!x' is a mere attitudinal failing, not a logical one.

In fact there are two issues here worth distinguishing. First, the inconsistency involved in the case of moral sentences and their negations needs to be inconsistency guaranteed by logical form. And this result is secured on the dominant commitment account: it is in virtue of their form that the attitudes 'α!β!x' and 'α!¬β!x' are inconsistent. Assuming that 'α!' is inconsistency-transmitting, they are inconsistent under any consistent interpretation of their terms.[11]

Second the account needs to explain the nature of the error made by agents who deny argumentative inferences whose validity depends on the inconsistency of moral sentences and their negations. In particular for the dominant commitment account, one might worry that having the attitudes 'α!β!x' and 'α!¬β!x' is a minor offence compared to having beliefs whose truth-conditions cannot simultaneously obtain. After all, it is not uncommon to wish for incompatible things, and perhaps *being for* incompatible things is no more heinous.[12] But again, the defender of the dominant commitment account can defer: she can claim that in both cases the mistake is having inconsistency-transmitting commitments directed at inconsistent contents. Further, in so far as the explanation as to why the α!-commitment is inconsistency-transmitting mirrors the explanation of why beliefs are inconsistency-transmitting, these errors will be equally serious. Of course, this reply relies on the availability of a uniform explanation of why beliefs and moral attitudes (such as the α!-commitment) are inconsistency-transmitting, but Schroeder's point applies here too: in the absence of grounds for pessimism about this explanation being available in both cases, it is reasonable for expressivists to proceed on the assumption that the explanation can be made to work in both sorts of case. (This possibility is explored in §3(c).) This proviso aside, I will henceforth assume that the dominant commitment account can explain the inconsistency of moral sentences and their negations in a way that sufficiently captures the argumentative force of avoiding such inconsistencies.

b. Generality Condition

Consider next the Generality Condition. So far the dominant commitment account has focused on simple moral sentences, but it easily generalizes to all simple sentences. According to this generalized view, 'x is Φ' expresses either an attitude or a belief. In both cases, '¬(x is Φ)' expresses a mental state describable using the same dominant commitment-operator (attitude or belief) only with a negation inserted immediately after that operator. In schematic terms, if the simple sentence 's' expresses the commitment Cψ then '¬s' expresses the commitment C¬ψ.[13] Here 'C' stands for a given commitment-type and 'ψ' stands for the content of that commitment. In the case of simple moral sentences, the commitment is an attitude (such as 'α!') whose content is an attitude directed at things (such as 'β!x'). In the case of simple non-moral sentences, the commitment is a belief whose content is a proposition. The semantic function of negation is univocal across these cases.

Unfortunately such an account runs into difficulties when we consider complex sentences that embed both moral and non-moral sentences, such as '¬(m or n)' where 'm' is a moral sentence and 'n' is a non-moral sentence. The resulting complex sentence is inconsistent with both 'm' and 'n'. An account of negation must explain this. For the dominant commitment account this means specifying the commitment expressed by the complex sentence. The problem is that whatever this commitment type is, it cannot be the same both as the dominant commitment involved in 'm' (which is an attitude) and as the dominant commitment involved in 'n' (which is a belief). This entails that the dominant commitment account's preferred method of explaining inconsistency – the A-type model – cannot apply in all cases.[14]

Is this a problem for the dominant attitude account? It need not be. There are at least two possible ways for the dominant attitude account to be developed so as to avoid this problem, both of which satisfy the Generality Condition.

The first approach is to abandon the A-type model for explaining inconsistency, at least in some cases. This is by no means a fatal admission. As we saw above, even the A-type model is incomplete as an explanation of sentential inconsistency until it can be explained why certain commitment types are inconsistency-transmitting. The possibility remains that the explanation of the

inconsistency of token commitments of the same type directed at inconsistent contents generalizes to a form that includes explanations of the inconsistency of token commitments of different types (even, perhaps, token commitments with consistent contents). At least, until the explanatory gap in A-type model is plugged, this possibility cannot be ruled out. On such view, the A-type model of explaining inconsistency is not redundant, but is a particular manifestation of a more general explanatory form. (Again, I believe that this possibility is realized, as I explain in §5(c).) For the time being therefore, note that abandoning the universality of the A-type model of inconsistency is not necessarily fatal, given that that model is itself incomplete.

The second response is more radical, yet preserves the A-type model in all cases. On this view, all assertoric sentences, moral and non-moral, simple and complex, express tokens of the same dominant commitment type and this commitment type is inconsistency-transmitting. This may be called the single dominant commitment account.

As Schroeder points out (2008a, 597), the single dominant commitment account is radical, but not without its attractions. It preserves the A-type model of explaining inconsistency in all cases, and this model has a good precedent in the case of belief. Obviously, it owes us an account of what the single dominant commitment type is. Expressivists are well-advised not to take this to be belief, since then their distinctive account of the meaning of moral sentences is lost. A more promising alternative is that the commitment is an attitude, perhaps the attitude of 'being for'. One startling consequence of this way of developing the view is that beliefs (the states expressed by some non-moral sentences) turn out to be a particular kind of non-cognitive attitude. Schroeder suggests that believing that x is F might be construed as *being for*: *proceeding as if x is F* (2008b, 93-5). Though such a move undermines many of the stated motivations for expressivism – which typically rely on a sharp distinction between belief and attitude – at least it provides a unified semantics of negation. Expressivists may hope that shifting the motivation for the theory is small price to pay for a semantics that works.

Still, the single dominant commitment view is a bold step for expressivists to take and many would think that if this only way for expressivists to solve the negation problem, then expressivism is in trouble. Schroeder has sympathies with this line of

thought: one way of understanding his geeky elucidation of the single dominant commitment view is as a *reductio* of expressivism.[15] My own view is that expressivists are best not to adopt single dominant commitment account: they should beware geeks bearing gifts. But, however that view develops it does at least preserve the univocality of negation (and the generality of A-type explanations). In other words, if the single dominant commitment account fails, it is not because it fails to meet the Generality Condition.

c. Fregean Condition

Consider finally the Fregean Condition. An account of negation must explain the fact that the meaning of sentences, both simple and complex, remains constant across negated and unnegated contexts. A well-known argument purports to show that no version of expressivism – including the dominant commitment account – can meet this condition, at least for any indirect context. The argument is the crux of the so-called Frege-Geach problem for expressivism and although the standard example concerns conditionals, a structurally isomorphic argument can be given for negation. The argument runs as follows.

According to expressivists, the meaning of 'x is M' is given by the attitude it expresses. This attitude is not expressed in the negated sentence '¬(x is M)'. Therefore, the meaning of 'x is M', as it appears in the negated context, cannot be the same as when it appears in the unnegated context. Hence the Fregean condition is not met.[16]

As a decisive refutation of expressivism, the Frege-Geach problem is too quick. There is in fact a well-known scheme of solution to the problem, originally suggested by Hare (1970). The problem is that adherents to the dominant commitment account cannot adopt this approach. I will first outline the Harean position and explain why it promises to solve the Frege-Geach problem, before showing why the dominant commitment account cannot adopt it.

Hare claimed that the attitude expressed by 'x is M' remains 'in the offing' in the attitude expressed by complex sentences such as '¬(x is M)'. This explains the constancy of meaning across these contexts. Cashing out the metaphor, expressivists

can hold that '¬(x is M)' expresses a complex mental state that is some psychological function of the mental state expressed by 'x is M', where the particular psychological function in play is isomorphically determined by the relevant sentential operator, in this case negation. Here 'psychological function' is construed narrowly so that a mental state or commitment is some function of another only if it takes the latter as one if its functional parts. Schroeder (2008a, 574) calls this position 'compositional semantics', since the meaning of complex sentences is given by complex compositional states whose functional parts are the mental states expressed by simple sentences. Compositional semantics seems a promising response to the Frege-Geach problem. It will explain the constancy of sentential meaning across simple and complex contexts in terms of the very same attitude being 'in the offing' in both contexts. In the present case, it is because the commitment expressed by '¬(x is M)' is some function of the commitment expressed by 'x is M' that the meaning of the sentence is constant across the contexts. Thus the Fregean Condition is met.

Unfortunately, the dominant attitude account cannot avail of this way of meeting the Fregean condition, because it isn't a version of compositional semantics. On the account, the sentence 'x is M' expresses an attitude of the form 'α!β!x'. Conversely, '¬(x is M)' expresses an attitude of the form 'α!¬β!x'. The problem is that the latter is not, in the relevant sense, a function of the former. Rather both are functions of a distinct attitude, 'β!x'. Hence the dominant commitment account cannot avail itself of the Harean solution to the Frege-Geach problem.[17]

The defender of the account might reply: there *is* an attitude that remains in the offing across negated and unnegated contexts, namely the attitude 'β!x'. This is the attitude that is α!-ed in the unnegated contexts and whose absence is α!-ed in the negated context. Perhaps the fact that this attitude is 'in the offing' in both contexts can explain the constancy of meaning across them. But though this may explain some constancy of meaning, it is the wrong type. The relevant explanandum is how 'x is M' has constant meaning across these contexts, and the meaning of 'x is M' is *not* given by the attitude 'β!x', but by the attitude 'α!β!x'. So although the dominant commitment account can deliver some constancy of meaning, it cannot deliver the constancy of meaning of 'x is M'. Thus the Frege-Geach problem is reinstated and the dominant commitment account fails to meet the Fregean Condition.[18] Expressivists must look for alternatives.

3. Commitment Semantics

a. Rejection

What is the mental state expressed by '¬(x is M)' where 'x is M' expresses a moral attitude? Let us follow the Harean suggestion that it is a state in which the original attitude remains 'in the offing'. The development of this view has recently been christened 'commitment-semantics' (Hale, 2002, 145). As Blackburn puts it

> the key idea here is that of a functional structure of commitments that is isomorphic with or mirrored by the propositional structure that we use to express them (Blackburn 1998, 71).[19]

The dominant commitment account doesn't achieve this isomorphism, because the propositional structure of '¬(x is M)' is of a dominant negation operator, but this was not mirrored in the psychology.

Let's follow Hare's footsteps more closely. Suppose the syntactic form of '¬(x is M)' is mirrored in the attitude it expresses, so that in describing this attitude negation *does* occur outside the attitude operator, as '¬α!x'. Of course, so understood '¬' must be given a different interpretation to the one at work in Indifference or Agnosticism. What might this be? An obvious answer is *rejection*.[20] To reject a moral attitude is to be committed to adopting an opposed attitude. Roughly it is to think that whatever the correct way to respond to the evaluated object is, it isn't like *that* (where 'that' refers to the rejected response). This thought is itself an attitude of response to the evaluated object: it is the attitude of someone who has ruled out responding to the evaluated object in the rejected way, but is open to responding to it in any other way.

This brief account of rejection can be made clearer by considering in more detail the type of commitment rejected. Modern expressivists hold that moral attitudes are settled practical stances or policies of action: to adopt a moral attitude towards an object is to have a policy of response to objects of that type, or to objects with certain of its features. As one expressivist puts it, the function of a moral attitude is to "mediate the move from features of the situation to a reaction" (Blackburn 1993, 168). To reject a moral attitude, therefore, is to reject a policy of reaction. It is to

adopt the policy that might be characterized as 'whatever the correct way to respond to the world is, it isn't like *that*'. We can understand this rejection as a conditional policy: it is the policy of adopting some specific policy other than the rejected policy. In terms of functions from features to reactions, the attitude rejected can be represented by the move from features of a situation to a particular reaction (disapproval or censure, say) whilst the attitude of rejection can be represented as the move from the same features of the situation to some other reaction.[21] One final illuminating way of representing these policies and their rejections is by considering the logical space of possible policies in response to an evaluated object (possible functions from objects to reactions). Where the rejected policy can be represented as one particular function from the object to a reaction, the policy of rejection can be represented as a disjunctive commitment to one of the complementary set of functions.

To give meat to these bones, consider an example. To think murder wrong is to have a certain policy towards murder: to avoid doing it oneself and to encourage others to do likewise, say. This can be represented as a function from murderous acts to these reactions. To reject this policy is to be committed to some *other* reaction to murder, that is, to be committed to responding towards murder in some way other than avoiding it and encouraging others to do likewise. In terms of function, this can be represented as the function from murderous acts to some other reaction. It is this policy that gets expression in the judgment that murder is permissible. Note that in this sense, judgments of permissibility express conditional policies (a policy of adopting one amongst a set of reactions) that agents can adopt without being committed to any particular reaction. This reflects the fact that it is possible to hold that it is not the case that murder is wrong, for example, without having a definite view on whether murder is right, supererogatory, or 'merely permissible' (that is, neither right nor wrong nor supererogatory). On the commitment semantic approach these last three judgments express particular policies of response towards murder, whereas the judgment that murder is permissible (that is, the judgment that it is not the case that murder is wrong) expresses the policy of being committed to one or other of these reactions.

a. Fregean Condition

How does this account of negation fare with respect to the three conditions? Consider first the Fregean condition. This is met because the attitude expressed by the unnegated sentence remains 'in the offing' in the attitude expressed by the negated sentence. This is to say that the attitude expressed by the negated sentence is some psychological function of the attitude expressed by the unnegated sentence: the function is rejection. When an agent asserts a negated moral sentence they express an attitude than can be defined in terms of the attitude that it rejects; the former attitude remains in the offing as the object of rejection.

b. Generality Condition

Does the account generalize from simple moral sentences to non-moral sentences and complex sentences? Does it meet the Generality Condition? To do this, the account of rejection needs to generalizable: just as one can reject the commitments expressed by moral sentences, one must be able to reject the commitments expressed by simple non-moral sentences (such as 'x is F') and by complex sentences (such as conditionals). I take these cases in turn.

Expressivists take most simple non-moral sentences to express beliefs.[22] To have a belief is to represent the world as being a certain way. For example, to believe that grass is green represents the world as being such that grass is green. To reject a belief is to reject a representation of the world: it is to think that whatever the world is like, it isn't like *that*. This rejection is itself a belief (just as rejecting a policy is itself a policy): it is the belief that the world is some way other than the rejected belief represents it as being.

What of the commitments expressed by complex sentences, such as conditionals? Given the general form of commitment semantics, the commitments expressed by complex sentences will have functional structure isomorphic with the sentences that express them. For example the commitment expressed by 'if p then q' can be functionally described thus: it is the commitment to have the commitment expressed by 'q' should one have the commitment expressed by 'p'. Blackburn (1998a, 71) calls this state one of being 'tied to a tree' of commitments. To reject such a complex commitment is to rule out being so tied. It is commitment to one of the set

of complex commitments incompatible (mutually unrealizable) with being tied in this way. Again, this can be represented in terms of a logical space, this time the space represents various ways of linking the two more basic commitments. The rejection can be represented in terms of the complement of the commitment rejected.

Note that this account of the state of rejection is *constructive* in the sense that for any commitment expressed by a sentence, be it a belief, attitude (policy), or tree-tying commitment, we have a recipe for generating an informative account of the commitment expressed by the negation. Moral sentences express attitudes characterizable as functions from features to reactions and their negations express attitudes characterizable in terms of complementary set of functions ('that's not the way to respond to murder'). Most non-moral sentences express beliefs characterizable by their representational content and their negations express beliefs characterizable in terms of the complementary content ('the world is not like that'). Complex sentences express tree-tying commitments characterizable as functions between more basic commitments and their negations express commitments characterizable in terms of the complementary set of functions. Thus commitment semantics does not, as Schroeder accuses some extant expressivist accounts, simply describe the sort of thing the negation-commitment would need to be *if* the negation problem is to be solved.[23] It provides an informative psycho-functional account of such commitments that easily generalizes.

c. Semantic condition

What, finally, of the Semantic Condition? An account of rejection that generalizes is worth nothing if we cannot explain why commitments and their rejections are inconsistent. So what is inconsistent about having a commitment and yet rejecting it?

A-type inconsistency is the appropriate model here, at least in some cases. For example, to reject the belief that grass is green is just to believe that grass is not green, and these beliefs are inconsistent, because grass cannot both be green and not green. Likewise, to reject a policy of response to murder is just to adopt the conditional policy of having some other response to murder, and these policies are inconsistent because they cannot both be enacted. An agent cannot both decide that disapproval, say, is the appropriate way to respond to murder, and also decide that some response

other than disapproval is the appropriate way to respond to murder. Similarly, a tree-tying commitment and its rejection are inconsistent because they commit agents to ways of being tied (or untied) that cannot simultaneously obtain. In each of these three cases the psychological inconsistency of the mental states is the immediate explanans of the sentential inconsistency of the sentences used to express them.

This explanation of inconsistency must face the challenge left outstanding for the dominant commitment account: explaining why these commitment types are inconsistency-transmitting in a way that secures logical, as opposed to mere attitudinal, inconsistency.

Once again, two issues can be distinguished. First, the inconsistency on show in the case of negation must be guaranteed by logical form. And this condition is met. It is in virtue of their psychological form that commitments and rejections are inconsistent. For example, if 'x is M' expresses some policy of response towards x then '¬(x is M)' express the rejection of this policy, that is, the policy of having some other response to x. These policies are inconsistent in virtue of their form (they are inconsistent under any consistent substitution for 'x' and 'M'). The same applies in the cases of rejecting beliefs and tree-tying commitments.

Second, the mistake involved in having inconsistent moral attitudes, or inconsistent tree-tying commitments, needs to be sufficiently serious to underwrite the logical force of inferences based on these inconsistencies. In particular, having inconsistent moral attitudes, or tree-tying commitments, needs to be as serious as having inconsistent beliefs. To meet this second challenge, it is helpful to look at the case of beliefs and then generalize. There is clearly something wrong with believing that mutually unrealizable states of affairs obtain. But what? One answer is to refer to the distinctive function of belief. Beliefs, it is often said, aim at truth, at accurate representation of the world. This is what explains why believing that x is F is distinct from wondering whether x is F: the former represents our understanding of the lay of the land and will affect out actions accordingly.[24] It is also explains why believing mutually unrealizable states of affairs to obtain is mistaken: for the land can only lay one way, so having both beliefs frustrates the aim of belief. Reference to the function of beliefs explains why belief is an inconsistency-transmitting commitment type.

Consider next the case of attitudes. Take the attitude of 'being for'. This attitude, like desire, aims at the realization of its object. To *be for* mutually incompatible goals is to be set for failure in this regard. This is why 'being for' is

inconsistency-transmitting: being for inconsistent contents frustrates the aim of this particular commitment type. This fills the gap in the dominant commitment account's explanation of inconsistency. It also explains the failure involved in such inconsistency in a way that mirrors the account given in the case of beliefs, by reference to the function of the commitment in question.

What about moral attitudes? The function of these attitudes is to regulate the move from features of the world to reactions to it, thus guiding action and laying a ground for intersubjective coordination. Moral discourse plays an integral role in this co-ordinating exercise: moral assertions are not mere sounding off, they are persuasive attempts to influence the attitudes and hence actions of others. The policies expressed in moral discourse are thereby pushed into a public arena, put forward to be discussed, disputed, tested, refined and ultimately publically adopted. According to expressivists, therefore, in asserting 'x is M' one is recommending a particular policy of response towards x. To also recommend a conflicting policy of response frustrates the coordinating purpose for which moral attitudes are formed and is inconsistent in just this sense. To take the simplest example, to assert that 'x is wrong' is to publically recommend a policy of response towards x that involves, at a minimum, not x-ing. To assert 'It is not the case that x is wrong' is to recommend a conditional policy of response towards x that is consistent with x-ing. To assert both is therefore to recommend conflicting courses of action, thereby frustrating the very purposes for which moral attitudes are formed and discussed. It is in this sense that an agent making these assertions is inconsistent.[25] As in the case of belief, it is by reference to the function of the commitment type in question that we can understand the force of the error involved in having inconsistent tokens of that type. The function changes, but the form of explanation is constant.

Finally, consider tree-tying commitments of the sort expressed by conditionals. The function of such commitments is to rule in and rule out certain combinations of the more basic commitments.[26] To be tied up in incompatible ways frustrates this function, since it leaves the agent with no stable inferential position with respect to the more basic commitments. Again, the explanation is structurally identical with that given in the case of beliefs.

In all these cases, the failure of an agent with inconsistent commitments is to frustrate the function of the commitments in question. More precisely, where there is inconsistency guaranteed by logical form, as in negation, the set of commitments the

agents possesses are such as to *guarantee* the frustration of the functions of the commitments of that set.[27] This account sees no distinction between the sort of mistake involved in having inconsistent beliefs and that involved in having inconsistent moral attitudes. There is therefore nothing 'mere' about attitudinal inconsistency.

Inconsistency, then, involves frustrating the function of the set of commitments in question, and logical inconsistency occurs when this frustration is guaranteed by the form of the commitments. It seems reasonable to assume that any agent who partakes in the practice of forming beliefs, or moral attitudes, or tree-tying commitments, at least implicitly endorses forming states with their respective functions. It follows that to bundle a set of commitments that guarantees the frustration of these functions is to render oneself unintelligible. We do not know what to make of an agent who believes both that x is F and that it is not the case that x is F, since she adopts a set of commitments that is guaranteed to frustrate the goal of having such states at all. In general, therefore, logical inconsistency as understood here entails just this sort of unintelligibility.[28]

One pleasing consequence of this view of logical inconsistency is that it generalizes from cases involving one commitment type to cases involving two or more commitment types. That is, it moves beyond A-type inconsistency while including it as a special case. (This was the result required if one was to resist the move to the single dominant commitment account in §2(b).) For example, the set of commitments expressed by the sentences 'p', 'if p then q' and '¬q' may be a belief, a tree-tying commitment and a rejection of a policy. Yet in so far as an agent who has this set of commitments necessarily frustrates the function of the tree-tying commitment (which is to rule in and out more basic commitments) she is inconsistent. In so far as this frustration is guaranteed by the form of his commitments, she renders herself logically inconsistent.

In summary, the commitment semantic account of negation can explain the sentential inconsistency of sentences with their negations in terms of the psychological inconsistency of commitments and their rejections. Such a view doesn't necessitate viewing attitudinal inconsistency as in any way less serious than inconsistency in belief. Ultimately, all types of psychological inconsistency can be understood by reference to the frustration of the function of the respective commitments, and in some cases (such as negation) this frustration will be guaranteed

by the form of the commitments, rendering the agent unintelligible. Hence the Semantic Condition is met.

d. Objections

The commitment semantic account of negation can meet the Fregean, Semantic and Generality conditions. It is therefore worthy of serious consideration as an expressivist account of negation. In order to elaborate the view further, and ward off potential objects, some further remarks are necessary.

First, it is worth noting that rejection is not, in the normal sense, a second-order commitment. It is not a commitment whose object is another commitment (so it is not best formalized as 'R!α!x' for example). To reject a commitment is to take a stand on the very same issue as was the topic of the original commitment, but a diametrically opposed one. So for example, to reject a policy of action towards murder is to adopt a policy towards murder (the policy of responding to murder in some other way than the rejected policy). Similarly, to reject a belief about the color of grass is to take a stand on the color of grass. Rejection is an opposed commitment, defined in terms of the commitment it opposes but occupying an area of the same logical space. This point nullifies a possible revival of the Unwin/Hale problem: for if rejection just is a negative second-order commitment then it would seem indistinguishable from Agnosticism, since the Agnostic too disapproves of the first-order commitment. The undesirable inference is blocked if rejection is understood as a position opposed to the rejected position, not simply directed at it. Both the agent who rejects a commitment and an Agnostic rules themselves out adopting the rejected commitment. But the Agnostic also rules himself out adopting any commitment in the relevant area, whereas the rejecter has a disjunctive commitment to adopting *some* commitment in that area (other than the commitment rejected, obviously). In the examples already given, the moral rejecter has a conditional policy of response to murder; the non-moral rejecter has a belief about the color of grass. The Agnostic has neither. Thus rejection and Agnosticism are distinct. (The possible conflation of rejection with Indifference is more easily dealt with, since the Indifferent agent does not, whereas the rejecting agent does, rule themselves out adopting the rejected commitment). In addition, rejection (and not Agnosticism or Indifference) will

typically be accompanied by discursive dispositions to oppose the rejected commitment. If you believe that grass is green and I believe that it is not green I will, other things being equal, try to persuade you of the falsity of your belief, oppose your expression of it, and otherwise dispute with you in open discussion. But discussion is not always open. Where the interlocutor is threatening, or otherwise holds power, one is sometimes best to forgo trying to change their mind. In such cases the opposed commitment – the rejection – remains, though it doesn't manifest itself in the normal discursive ways.[29] Rather it persists as a view of the way the world is (in the case of belief) or of how to respond to the world (in the case of moral attitudes) or of how to combine more basic commitments (in the case of tree-tying commitments).

Second, the account of rejection offered here builds upon brief remarks about negation made by Blackburn. He notes one difference between rejection and agnosticism (which here includes both Indifference and Agnosticism), namely their

> relative 'robustness': it can take more to shift a definite attitude one way or another than it takes to shift agnosticism (Blackburn 2002, 168).

This feature of rejection is explained on the above account, for the rejecter of a belief accepts that the world is some way other than the rejected belief represents it as being; the rejecter of a policy holds that the way to respond to murder is some way other than the rejected policy suggests; and the rejecter of a tree-tying commitment holds that the way to be tied to combinations of basic commitments is some way other than that encapsulated in the rejected commitment. In the normal case the rejecter will have reasons for thinking that the world is this way, or that this policy is appropriate, or that this way of being tied (or untied) is justified; reasons which the Indifferent and Agnostic agents lack. To change the position of the rejecter requires overcoming or undermining these reasons, whereas changing the position of the Indifferent or Agnostic agent does not, hence the difference in relative robustness.

Finally, rejecting a commitment iterates in the same way as negation. 'x is Φ' is logically equivalent to '¬¬(x is Φ)'. If rejection is the psychological correlate of negation, this result must be mirrored in the psychology.[30] And it is. To reject a commitment is to take up the opposed position. To reject this rejection is therefore to

take up a position opposed to the opposition, which just is the original commitment. This is mostly easily seen in the case of belief. To reject a belief is to have a belief: it is to believe that the world is not like the original belief represents it as being. To reject this rejection is therefore to believe that it is not the case that the world is not like the original belief represents it as being, which is just to believe the world *is* as the original belief represents it as being. The case of policies (such as moral attitudes) is similar. To reject a policy is to have the policy of responding to the world in some way other than that given by the original policy. To reject this rejection is therefore to have the policy of responding to the world in some way other than *this* policy, which is to have the policy of responding to the world in the original way. Thus to reject the rejection of a policy is just to have the original policy.

4. The Expression Account

Commitment semantics is the most popular way for expressivists to tackle indirect contexts, so it is reasonable to attempt to apply it to the case of negation. But are there any alternatives? Consider the 'expression account'.

Return to the Hale/Unwin problem. Take the schema

(1') 'x is M' expresses ¬α!x

And then, moving to the negated case, the schema

(3'') '¬(x is M)' expresses ????

The problem is that in filling in the blanks in (3'') there is simply not enough complexity in the original schema; not enough places for negation. The dominant commitment account solves the problem by introducing some complexity inside the scope of the dominant commitment operator, that is, by replacing the second 'x' in (1') with 'β!x'. In effect, the commitment semantic account supposes that complexity can occur before the dominant attitude operator, so that 'x is M' expresses '¬α!x' (where the external negation stands for *rejection* as distinct from Indifference or Agnosticism). But another place for complexity is *within* the relation between the utterance and attitude (the expression relation).

To follow this suggestion, 'expression' needs unpacking. One account is that expressing a mental state involves, as a necessary condition, advertising an intention to defend that state in open discussion, other things being equal.[31] Suppose we substitute this into the original schema, which then becomes:

(1'') 'x is M' advertises an intention to defend α!x

This provides obvious scope for the requisite complexity. The external negation schema would be:

(3''') '¬x is M' advertises an intention to <u>attack</u> α!x

In the former case the agent advertises the intention to defend a particular way of responding to the world, the way encapsulated in the attitude 'α!x'. This defense will involve responding in that way oneself, but also the disposition to urge that mode of response upon others, by whatever argumentative means one considers appropriate. Agreement with the assertion will consist in coming to defend the same policy of response. In the latter case – the negation case – the agent advertises an intention to attack a particular way of responding to the world. This attack will involve not responding in that way oneself but also the disposition to urge people against that mode of response by whatever argumentative means one considers appropriate. Agreement with the assertion in this case will consist in coming to share the same propensity to attack.[32]

On this view, the Hale/Unwin problem is solved. The externally negated sentence has distinct meaning from the internally negated sentence, which advertises an intention to defend α!¬x. External negation will also be distinct from Indifference, which involves no offensive or defensive intentions with respect to either attitude. The Agnostic does possess some discursive intentions but not the one's involved in negation. The Agnostic's 'harrumphing' advertises an intention to attack, not the particular attitude expressed by an unnegated sentence, but any argument or consideration that could be offered in support of this attitude. In other words, what the Agnostic opposes is not the original unnegated sentence, but the assertion that any putative consideration provides reason to accept this sentence. The Agnostic does not possess an intention to attack the attitude expressed by the unnegated sentence, for to have this intention is to urge that the attitude encapsulates an inappropriate pattern of response to the world, and the Agnostic is precisely someone who denies that such judgments can be made. Thus the position of those who accept the negated sentence is distinct from Agnosticism.

Does this account of negation meet the three conditions? The Fregean Condition is met, although this is not a version of compositional semantics. Negation is not understood as involving a distinct functionally related attitude, but rather as involving a distinct semantic relation, related to yet distinct from the relation of expression. Despite shunning the Harean maneuver, however, the expression account can make some sense of the constancy of meaning. For the attitude expressed by 'x is M' is clearly in view in both unnegated and negated contexts: there is an intention to

defend it in (1'') and an intention to attack it in (3'''). This sameness of argumentative focus is a plausible explanans for the sameness of meaning across the contexts.

The expression account meets the Semantic Condition because the intentions involved in the two contexts are inconsistent: one cannot consistently intend both to attack and defend an attitude. Because the sentences advertise inconsistent intentions, they are themselves inconsistent. This is another example of the A-type model of explaining inconsistency.[33] Furthermore this inconsistency is obviously a result of the contribution of sentential negation to the negated sentence, for the function of the operator is precisely to switch the intention from a defensive to offensive mode (mirroring the reversal in truth-value that is the natural view of the role of negation), and it is this that generates the psychological inconsistency.

Finally, the Generality Condition is met because the mental state which one is advertising an intention to defend or attack could be a belief, attitude or tree-tying commitment. In all cases, negation functions to reverse the direction of the intention towards the commitment. The nature of the commitment itself is unimportant.

The Expression account meets the Fregean, Semantic and Generality conditions. Like the commitment semantics account, therefore, it deserves to be considered as a potentially viable expressivist account of negation.

5. Diagnosis of Schroeder's Error

Besides generously elucidating the dominant commitment account and offering it to expressivists as a way to develop their views, Schroeder argues that it is the only possible way for expressivists to solve the negation problem. The existence of two viable alternatives proves this conclusion false. So where might Schroeder's argument go wrong?

Schroeder holds that an account of the semantics of negation needs to do at least two things. First, to solve the Unwin/Hale problem we need an account that has the requisite complexity. The problem, remember, is that the schema (1') is too simple – it has too few places for negation. To solve the problem we need more complexity in this schema (2008a, 589-591). Second, Schroeder claims, the resulting account must explain the inconsistency of sentences and their negations, for which A-type inconsistency is a good model (2008a, 579).

Schroeder's mistake is assuming that only place for complexity to occur is within the scope of a dominant commitment operator, thus preserving the A-type model in all cases. In fact, there are two further places for the complexity to occur: outside the dominant attitude operator (as in commitment semantics) and within the relation between the utterance and the attitude (as in the expression account). Both views preserve the A-type model, at least in some cases. Furthermore, where that model is given up, the resulting inconsistency is not, as Schroeder claims, necessarily mysterious (Schroder, 2008a, 581) for it can be explained in terms of the frustration of the function of the commitments in question (which is, in any case, one way of plugging the gap in the A-type model preferred by the dominant commitment account). Schroeder's mistake, therefore, is to misidentify the number of places where the requisite complexity may occur.

It is worth remarking here on another advantage of the two alternatives to Schroeder's proffered dominant commitment account. Both commitment semantics and the expression account are not committed to any particular view of the nature of moral attitudes, whereas the dominant commitment account must hold that moral attitudes are ascended (thus having two places for negation within the dominant commitment). This can be considered an advantage for commitment semantics and the expression account, for two reasons. First, Schroeder's particular choice of the ascended attitude expressed by judgments of wrongness is 'being for: blaming for',

but this seems to unfairly preclude the possibility of substantive moral positions according to which wrongdoing and blameworthiness come apart, such as Smart's extreme utilitarianism (1956, 347). Within the framework of the dominant commitment account, it is not clear that there is any better candidate for the attitude expressed by judgments of wrongdoing that doesn't have the unhappy result of ruling out substantively interesting moral positions by conceptual fiat. Second, the dominant commitment account of negation cannot apply to sentences taken to be expressive of non-ascended attitudes, for they do not have enough places for negation. Yet some of these sentences – aesthetic sentences, for example – can be intelligibly negated. These points provide additional reasons for expressivists to politely decline Schroeder's gift.

6. Conclusion

The aim of this paper is not to provide a 'logic for the attitudes' that can explain all the logical relations between arbitrary sets of and moral and non-moral sentences.[34] Rather it is just to show that there are expressivist accounts of negation that can explain its logical properties and that generalize. Such accounts are worthy of further consideration. Yet these accounts are not totally detached from the traditional understanding of logic with which the argument began. On one understanding at least, there is nothing more to a set of sentences being systematizable in truth-conditional terms than those sentences being able to function in the normal ways in logical contexts such as negation.[35] On this view to show that sentences expressive of attitude can function in negation is to go part of the way of earning the right for those sentences to be systematized in truth-conditional terms. Understood this way, the expressivist accounts of negation offered here are not alternatives to the truth-conditional understanding of negation with which we began, rather they are attempts to earn the right to the sort of systematization that talk of truth-conditions provides. If the coin is good, there is no attitudinal logic distinct from standard logic, just a deeper, perhaps surprising, route to the latter. The journey may be longer, but the destination reassuringly familiar.

References

Barker, S. (2006) "Truth and the Expressing in Expressivism," *Metaethics after Moore*, eds. T. Horgan and M. Timmons. Oxford: Clarendon Press

Blackburn, S. (1984) *Spreading the Word*. Oxford: Clarendon Press.

——. (1988) "Attitudes and Contents", *Ethics* 98, 501-17.

——. (1993) *Essays in Quasi-Realism*. Oxford: Oxford University Press.

——. (1998) *Ruling Passions*. Oxford: Clarendon Press.

——. (2001) "Reply," *Philosophical Books* 42, 27-32.

——. (2002) "Replies," *Philosophy and Phenomenological Research* 65, 164-176.

——. (2006) "Antirealist Expressivism and Quasi-Realism," *The Oxford Handbook of Ethical Theory*, ed. D. Copp. Oxford: Oxford University Press.

Björnsson, G. (2001) "Why Emotivists Love Inconsistency," *Philosophical Studies* 104, 81-108.

Dreier, J. (2006) "Negation for Expressivists: A Collection of Problems with a Suggestion for their Solution," *Oxford Studies in Metaethics, volume 1*. Oxford: Clarendon Press.

Elstein, D. (2007) "Against Sonderholm: Still Committed to Expressivism," *Proceedings of Aristotelian Society* 107, 111-116.

Geach, P. (1965) "Assertion," *Philosophical Review* 74, 449-65.

Gibbard, A. (1990) *Wise Choices, Apt Feelings: A Theory of Normative Judgment*. Oxford: Oxford University Press.

——. (2003) *Thinking How to Live*. Cambridge, Ma.: Harvard University Press.

Hale, B. (1993) "Can there be a logic of the attitudes?" *Reality, Representation and Projection*, eds. J. Haldane and C. Wright. Oxford: Oxford University Press.

——. (2002) "Can Arboreal Knotwork Help Blackburn Out of Frege's Abyss?" *Philosophy and Phenomenological Research* 65, 144-49.

Hare, R.M. (1970) "Meaning and Speech Acts," *The Philosophical Review* 79, 3-24.

Horgan, T, and Timmons, M. (2006) "Cognitivist Expressivism," *Metaethics after Moore*, eds. T. Horgan and M. Timmons. Oxford: Clarendon Press

Schroeder, M. (2008a) "How Expressivists Can and Should Solve Their Problem with Negation," *Noûs* 42, 573-599.

——. (2008b) *Being For: Evaluating the Semantic Programme of Expressivism*.

Oxford: Oxford University Press.

Schueler, G.F. (1988) "Modus Ponens and Moral Realism," *Ethics* 98, 501-17.

Searle, J. (1962) "Meaning and Speech Acts," *The Philosophical Review* 71, 423-32.

Sinclair, N. (2008) "Free Thinking for Expressivists," *Philosophical Papers* 37, 261-286.

Smart, J.J.C. (1956) "Extreme and Restricted Utilitarianism," *Philosophical Quarterly* 6, 344-54.

Stevenson, C.L. (1937) "The Emotive Meaning of Ethical Terms," *Mind* 46, 14-31.

Unwin, N. (1999) "Quasi-Realism, Negation and the Frege-Geach Problem," *Philosophical Quarterly* 49, 337-352.

Velleman, D. (2000) "On the Aim of Belief," *The Possibility of Practical Reason*. Oxford: Oxford University Press.

Wright, C. (1987) "Realism, Antirealism, Irrealism, Quasi-realism," *Midwest Studies in Philosophy* XII, 25-49.

——. (1992) *Truth and Objectivity*. Harvard: Harvard University Press.

[1] The two alternatives are distinct from those offered by Gibbard (2003) and Horgan and Timmons (2006). For problems with the former see Dreier (2006) and with the latter see Schroeder (2008a, 582-4).

[2] See e.g. Gibbard (1990, 2003) and Stevenson (1937).

[3] Note that 'simple' and 'complex' as defined here are not exhaustive. Some sentences, such as questions and commands, are neither.

[4] Though I don't have the space to elaborate here, these conditions generalise to other instances of the embedding problem.

[5] For this point see Unwin (1999, 343) and Schroeder (2008a, 575).

[6] As Unwin (1999, 352) recognises.

[7] This suggestion is made by Unwin (1999, 349-52) and pursued at length by Schroeder (2008a, 2008b).

[8] See Blackburn (1998a, 8-14) and Gibbard (1990, chap. 3).

[9] Initially, Schroeder (2008a, 577) seems to suggest that two contents are inconsistent just when they cannot both be true. The account offered here extends this idea to cover cases where contents do not have truth conditions, such as 'blaming for murder'.

[10] See Hale (1993, 339) and Wright (1987, 33).

[11] See Schroeder (2008a, 594).

[12] See Schueler (1988).

[13] See Schroeder (2008a, 591-2).

[14] See Schroeder (2008a, 596-7).

[15] See Schroeder (2008b, 14). I should add that I don't consider 'geeky' to be derogatory; it seems to me to be synonymous with 'rigorous' and 'scholarly'.

[16] See Geach (1965). Searle (1962) gives a structurally similar argument in terms of speech acts: since there is no common speech act across the negated and unnegated contexts, a speech-act theory of meaning cannot preserve constancy of meaning across them.

[17] Diagnosis: Compositional semantics introduces is an isomorphism between the logical form of sentences and the psycho-functional form of the *commitment* they express. The dominant commitment account introduces an isomorphism between between the logical form of the sentences and *the contents of* the dominant commitment. The slide between the two is a subtle but important one in Schroeder (for the former see 2008a, 574, for the latter see 592-3).

[18] Note that at this point a similar objection might apply to the case of non-moral sentences. Where F is a non-moral predicate, 'x is F' expresses a belief that x is F, '¬(x is F)' expresses a belief that ¬(x is F), yet the latter is not a function of the former. This would seem to demonstrate that all expressivists face a Frege-Geach problem for their account of sentences expressive of belief, and even more worryingly, a problem which the Harean manoeuvre cannot avoid. However, the problem can be avoided so long as the belief that ¬(x is F) can be understood as some function of the belief that x is F. Commitment semantics, discussed below, secures this result.

[19] It doesn't follow that the functional structure can always be identified independently of the logical articulation; see Blackburn (2002, 166-7). For developments of commitment semantics, see Blackburn (1988, 1998, 2002), Bjornsson (2001), Elstein (2007), Hale (1993) and Sinclair (2008).

[20] See Blackburn (1988, 192 and 2002, 167). Bjornsson (2001, 88 & 94) talks of 'negative opinions'.

[21] Formally, an attitude, α!x, can be represented as a policy function from stimulus S to reactions R, thus: Pol[R, S]. The rejection, or ¬α!x, can be represented as Pol[¬R, S], where '¬R' signifies some response other than R, thus 'driving the negation inwards' (Blackburn, 1988, 192). The particular formalism chosen is less important than the philosophical understanding of the semantic relations that it formalises. Note that, in describing policies the brackets shouldn't be thought to signal ascent (this is what distinguishes the view from the dominant commitment account). The formalism doesn't signify a distinct higher-order policy directed at a reaction/stimulus pair; rather the policy is characterizable by this pair.

[22] The exceptions are non-moral sentences that express attitudes. Perhaps aesthetic sentences are like this. The account of rejection in their case is identical to the account of rejection in the case of moral attitudes, since nothing in that account relied on a particular feature of the moral reaction.

[23] See Schroeder (2008a, 587-2), Dreier (2006, 22) and Bjornsson (2001, note 7).

[24] See Velleman (2000).

[25] The inconsistency is more complex in the case of evaluative concepts, such as 'good', whose connection to practical guidance is less direct. But so long as the function of a moral concept is to provide practical guidance of some kind, inconsistent application of that concept will consist in recommending inconsistent courses of action.

[26] Why might one form states that function this way? To keep track of the implications and consistency relations between more basic commitments. See Blackburn (1998, 71).

[27] A weaker type of inconsistency is involved where the functions of the commitments are only contingently frustrated, such as when an agent desires outcomes that only happen to be mutually unrealizable. It is the goal of enquiry to uncover such inconsistencies.

[28] Note how this account is consonant with Blackburn's account of illogicality as unintelligibility. See Blackburn (1998, 72 and 2002, 167).

[29] This is a deficiency in the account of rejection given by Sinclair (2008, 267).

[30] As Elstein (2007) realises.

[31] See Blackburn (2001, 2006) and Barker (2006, 304-5). Note that Barker himself cannot accept the account of negation on offer here, since he holds that all expressive assertion advertises the intention to defend some commitment or other.

[32] Note that commitment semanticists can accept the analysis of expression suggested by the expression account, but will reject the claim that negation involves a distinct kind of relation between utterances and commitments.

[33] See Schroeder (2008a, 577).

[34] See Blackburn (1988) for one such attempt. Both accounts of negation offered here are consistent with, but do not require, the formal model of consistency Blackburn offers in this paper.

[35] I refer to various 'minimalist' views of truth-aptness. See Wright (1992).