Philosophy Compass 3/4 (2008): 703–720, 10.1111/j.1747-9991.2008.00155.x

What is the Frege-Geach Problem?

Mark Schroeder*
University of Southern California

Abstract

In the 1960s, Peter Geach and John Searle independently posed an important objection to the wide class of 'noncognitivist' metaethical views that had at that time been dominant and widely defended for a quarter of a century. The problems raised by that objection have come to be known in the literature as the *Frege-Geach Problem*, because of Geach's attribution of the objection to Frege's distinction between content and assertoric force, and the problem has since occupied a great deal of the attention both of defenders of broadly noncognitivist views, and of their critics. In this article I explain Geach and Searle's historical objections, and put the subsequent discussion into dialectical context, paying some attention to the developments along the way and how they have enhanced our overall understanding of the problem. The article covers a lot of territory, so we will only be able to see the highlights, along the way. For further reading, see the Works Cited.

What is Noncognitivism?

Classificatory labels like 'emotivism', 'noncognitivism', and 'expressivism' have been used in a variety of ways. For the purposes of this article, I will use 'noncognitivism' as a catch-all label for the wide and heterogeneous class of views which includes Ayer, Carnap, Stevenson, Hare, Blackburn, and Gibbard as primary exemplars. These authors differ widely over how they believe moral language to work, but they all agree that the *kind* of meaning that moral terms (like 'wrong') have is importantly different from the *kind* of meaning that descriptive terms (like 'green') have.

Among emotivist views, which were the first kind of noncognitivist views defended, though all shared the view that moral terms have a different kind of meaning than descriptive terms, and all agreed that moral terms had something to do with the emotions, there was a fair bit of disagreement over just what moral terms did have to do with the emotions. According to some views, for example, moral sentences are used to *create an effect* – to elicit an emotion on the part of the audience. While according to other views, moral sentences are used to *express* or *give voice to* the emotions of the *speaker*. Sometimes these views were put together, either explicitly or without discussion, and sometimes they were put together with other views about the use of moral language, which didn't directly

relate to the emotions, strictly speaking. For example, some theorists said that moral sentences are disguised imperatives, or even that they are disguised commands.

When R. M. Hare published *The Language of Morals* in 1952, he was highly critical of earlier noncognitivists for assimilating moral language to other kinds of language. Moral language, Hare held, is neither *just like* exclamations, as Ayer's colorful prose suggests, nor *just like* commands, as Carnap insisted. Instead, Hare held, what was right about the noncognitivist family of views was that moral language belongs to a broad family of language, *prescriptive* language, of which imperatives are another and more familiar instance. Moral sentences are not disguised imperatives, on Hare's view, but they do have a meaning of broadly the same kind as imperatives, contrasting with the kind of meaning that ordinary descriptive sentences like 'grass is green' have.

Contemporary versions of noncognitivism differ from each of these previous classes of views in yet another way. Over the last quarter century, Simon Blackburn, Allan Gibbard, and Terry Horgan and Mark Timmons have developed views which fit into the tradition descending from Ayer and Stevenson, but which have a more detailed flavor that has come to be known as *expressivism*.[1] The basic idea of expressivism is that it is the job of a semantic theory to explain what a sentence, 'P', means, by saying what it is to think that P. Whatever a given theory says it is to think that P, that is the mental state that the theory counts as being *expressed* by 'P'. So a complete expressivist theory would assign each sentence of the language to a mental state which it expresses, and this would constitute, according to expressivists, a semantic theory for the language. Metaethical expressivists like Blackburn, Gibbard, and Horgan and Timmons go on, further, to explain that moral sentences differ in the kind of meaning that they have from ordinary descriptive sentences, because the kind of mental state that they express differs from the kind of mental state expressed by ordinary descriptive sentences.

So noncognitivist views differ widely both in their commitments and in their theoretical framework. The safest way to characterize what all of these views have in common, is that they reject the idea that moral terms have the same kind of meaning as ordinary descriptive terms. Once we characterize noncognitivist views in this way, moreover, it is easy to characterize the crux of the Frege-Geach Problem. It is that there is no linguistic evidence whatsoever that the meaning of moral terms works differently than that of ordinary descriptive terms. On the contrary, everything that you can do syntactically with a descriptive predicate like 'green', you can do with a moral predicate like 'wrong', and when you do those things, they have the *same semantic effects*.

So the Frege-Geach Problem is at bottom the problem of how it could be that moral and descriptive terms have exactly the same sort of semantic properties in complex sentences, even though they have different kinds of

meaning. In the following sections I'll trace the historical development of this very general problem, of which a relatively narrow set of issues has occupied philosophers' attention.

Geach and Searle's Original Objection

Geach and Searle's objections were originally formulated against noncognitivist views which they understood to involve a certain kind of claim about the *speech acts* involved in asserting moral sentences. Hare had said that to call something 'good' is to *commend* it, and took this to tell us something about the *meaning* of 'good' – not just a fact about what people in general happen to use the word 'good' in order to do. So Geach took Hare to be committed to the view that a given appearance of the word 'good' means *good* just in case it is being used to commend. That is why he chose examples in which it is clear that the word 'good' is *not* being used to commend. Some of the best examples of this were questions, negations, and the antecedents of conditionals: in ordinary cases, people do not say 'is this good?' in order to commend it, nor do they say, 'this is not good' or 'if this is good, then that is good' to commend it. So, Geach inferred, it follows from Hare's commitments that 'good' must mean something else when it appears in these sentences, because it is not being used to commend.

Geach's argument did not end there, however; instead, he argued that 'good' *must* mean the same thing in these sentences as it does in 'this is good'. That is because 'this is good' is the *answer* to 'is this good?', because 'this is good' *contradicts* 'this is not good', and because 'that is good' follows logically, by *modus ponens*, from 'this is good' and 'if this is good, then that is good'. What Geach was doing, in offering this argument, was showing that these *semantic properties* of questions, negations, and conditionals – of what their answers are, of what contradicts them, and of what logically valid arguments they figure in – are explained by the fact that the terms involved *mean the same thing* as they do in the unembedded sentence, 'this is good'. So Geach concluded that Hare's view was in a very bad way. It was committed to denying exactly what was necessary in order to explain the semantic properties of these complex sentences.

Searle's objection was essentially the same, although he offered an extra twist. Searle offered the same kind of evidence as Geach that moral terms must have the same meaning when embedded as when unembedded, but he offered the noncognitivist a more sophisticated position, according to which 'good' doesn't have to *always* be used to commend, in order to have its usual meaning, so long as it is *connected* to commendation in the right sort of way. It's a little bit unclear exactly what this right sort of way was supposed to amount to, but we can get a hint from looking at Searle's own view about 'promise'.

Searle held that we can understand the meaning of 'I promise to do it' by understanding that it is used to *promise* – a view that sounds very much like Hare's view that we can understand 'this is good' by understanding that it is used to *commend*. So unlike Geach, Searle had to be careful in order to distinguish his own view, which he thought was defensible, from Hare's, which he thought was not. I think the answer is that Searle thought that 'I don't promise to do it' means, 'I don't perform the speech act of promising to do it', whereas 'this isn't good' *doesn't* mean, 'I don't perform the speech act of commending it'. So for Searle, the meaning of 'promise' *can* be understood *somehow* in terms of the speech act of promising, even in complex sentences, but the meaning of 'good' *can't* be understood in terms of the speech act of commending in such complex sentences – or at least, not in the same way.

Whatever the complications in Searle's version of the objection, however, both Geach and Searle were pressing the objection that noncognitivists are committed to denying that 'good' or 'wrong' mean the same thing in at least certain kinds of embedded contexts as they do in simple atomic sentences, and that this is bad, because we *need* to assume that they mean the same thing in both places, in order to explain the semantic properties of the complex sentences.

Hare's Answer: Compositional Semantics

Hare replied to Searle in 1970, and gave the answer to this objection that has informed essentially all substantive approaches to the Frege-Geach Problem since then. Hare argued that the problem noncognitivists face in accounting for the meaning of complex sentences is essentially no different from the problem that everyone faces in accounting for the meaning of complex ordinary descriptive sentences. For example, 'this is not green' does not have the same truth-conditions as 'this is green', but that doesn't stop ordinary truth-conditional semanticists from holding that 'green' means the same thing in both sentences. So why should the fact that 'this is not good' is not used to perform the same speech act as 'this is good' mean that noncognitivists are forced to hold that 'good' doesn't mean the same thing in both sentences?

The answer offered by ordinary truth-conditional semanticists, as Hare understood it, is to say that though the truth-conditions of 'this is not green' are not the *same* as those for 'this is green', they are still a *function* of those truth-conditions – a function given by the meaning of the word 'not'. So 'this is green' means the same thing in both sentences, because it makes the same *contribution* toward the truth-conditions of the whole sentence. So similarly, Hare held, all that the noncognitivist needs in order to be able to say the same thing, is to hold that the speech act performed by 'this is not good' – which is, of course, something other than commendation of the referent of 'this' – is a *function* of the speech act

performed by 'this is good' – a function given by the meaning of 'not'. Then, Hare will say, 'good' means the same thing in both sentences, because it makes the same contribution to the speech act performed by the whole sentence.

There is an important hitch for Hare's answer to Geach and Searle, which it is important to appreciate before we go on. Strictly speaking, the reason why ordinary truth-conditional semanticists can say that the truth-conditions of 'this is not green' are a function of the truth-conditions of 'this is green', is that they think that 'this is green' *really has* truth-conditions, even when it is embedded in the larger sentence, 'this is not green'. But Geach's point was not just that someone who says 'this is not good' does not use the *whole sentence* to commend the referent of 'this'; it was that in ordinary cases someone who says 'this is not good' does not engage in any speech act of commendation *at all* – even by a proper *part* of the sentence.

If Hare wants to solve this problem, he needs to associate the meaning of a moral sentence with a speech act that it is *suited for* – not with one that is actually performed by each occurrence of a sentence with that meaning. This should not be a surprise; just as speakers can utter complex sentences which contain 'this is good' as parts without commending the referent of 'this', they can also utter 'this is good' sarcastically or in other ways that do not involve commending the referent of 'this', either. Similarly, however, speakers can utter imperatives without issuing commands. For example, they can be uttered in jest, or to convey information. So Hare should say that the relationship between 'this is good' and commendation is like the relationship between 'do this' and commanding, and say that it is the job of a semantic theory to assign each sentence to the speech act that it is in this sense *suited* to perform, assigning suitable speech acts to complex sentences as a function of the speech acts assigned to their parts that is given by the meaning of the words – like 'not' and 'if . . . then' that are used to construct the complex sentence.

Hare's answer to Searle therefore meets the terms of Geach and Searle's original challenge: it explains how moral terms like 'good' could have the same meaning in both places in which they appear. But importantly, it doesn't yet solve the problem raised by their arguments, because not just any function assigning a speech act to be used by 'this is not good' will suffice to explain why it has the semantic properties that it does. Just to take two obvious examples, neither the identity function nor the function which maps every speech act to the speech act of promising to get married before July would yield an adequate semantic account of the meaning of 'not'.

The New Shape of the Problem

So if noncognitivists take Hare's answer to Geach and Searle seriously, they still owe us, for every complex-sentence-forming construction in natural languages, an account of just *what* gets assigned to the complex

sentence as a function of the assignment to its parts, and an explanation of why this semantic theory yields the right predictions about the semantic properties of questions, negations, conditionals, and so on. Speech-act theories like Hare's will need a compositional semantics which assigns to every sentence the speech act that it is suited for performing, and contemporary expressivist views will need a compositional semantics which assigns to every sentence the mental state that it expresses. So for different kinds of noncognitivist view, their semantic theory will take different forms, but these views can still follow the broad outlines of Hare's suggestion.

This is what 'solutions' to the Frege-Geach Problem are really trying to do. They are trying to fulfill Hare's promise that a noncognitivist view can do the same thing as an ordinary truth-conditional view, and provide a compositional semantics for at least certain linguistic constructions which tells us the meaning of complex sentences of a certain kind in terms of the meanings of the parts of that sentence – either in terms of speech acts or more commonly, in terms of mental states expressed – and then tries to show that this is an adequate semantics for the sentence, because it can predict and explain the sentence's semantic properties. For example, an adequate semantics for 'not' must explain why negated sentences contradict the sentences they negate, and an adequate semantics for conditionals must explain why they license *modus ponens*.

The problem is *very big*, because for every complex-sentence-forming construction in natural languages, sentences formed using that construction using moral terms like 'good' have the same sort of semantic properties as sentences formed using that construction using ordinary descriptive terms like 'green'. This is true not only for questions, negations, and conditionals, but also for quantifiers, modals, tense, attitude-verbs, generics, adverbs of quantification, intensifying adverbs like 'very', and so on. Noncognitivists believe that moral terms have a different kind of semantics than ordinary descriptive terms, but somehow every complex-sentence-forming construction manages to do exactly the same sort of things with them that it does with ordinary descriptive terms.

This is the new shape of the Frege-Geach Problem, and it is the one that noncognitivists have been trying to address since Hare. The problem is to construct a compositional semantics for natural languages which makes complex moral sentences and complex descriptive sentences turn out to have the same kinds of semantic properties – and the *right* kind of semantic properties – even though moral and descriptive terms really have two quite different kinds of meaning.

Early Approaches to Conditionals: Higher-Order Attitudes

Despite the broad scope of the Frege-Geach Problem, most research devoted to it during the 1980s and 1990s focused on the case of conditionals, and specifically of explaining why *modus ponens* is a valid rule of inference.

Since the vast majority of the work in this period focused on the case of expressivism, I'll now restrict our attention to such views – although very similar issues would arise for a speech-act semantics like Hare's. Expressivist treatments of 'wrong' generally hold that atomic 'wrong' sentences express some negative attitude toward their subjects. Without loss of generality, I'll call this negative attitude *disapproval*, even though different expressivist views will have different theories about precisely what this attitude involves and what it should be called. Schematically, then, such expressivist views hold that for any value of 'X', 'X is wrong' expresses a negative attitude called *disapproval* toward the referent of 'X'.

An important early category of expressivist approaches to conditionals treated them as expressing *higher-order attitudes* toward the attitudes expressed by their parts. So, for example, Simon Blackburn (*Spreading the Word*) proposed, in a development of an idea implicit in his 1973 paper, 'Moral Realism', that 'if stealing is wrong, then murder is wrong' expresses disapproval of the state of both disapproving of stealing and not disapproving of murder. So it expresses a higher-order attitude toward the mental states expressed by the parts of the sentence.

Blackburn's approach is designed to explain why someone who accepts 'stealing is wrong' and 'if stealing is wrong, then murder is wrong' is under a kind of rational pressure to accept 'murder is wrong' – or at least to give up on one of the other two.[2] This is because so long as she accepts 'stealing is wrong' and *doesn't* accept 'murder is wrong', she is in the very state that she disapproves of, in virtue of accepting the conditional. So Blackburn held that there is a kind of 'incoherence' in her attitudes – an incoherence that can be resolved by going on to accept the conclusion, or by giving up on one of the premises.

Many complications and problems for accounts like this one were pointed out between the mid-1980s and the mid-1990s, and I can't survey all of them here. But in his 1996 article, 'Expressivism and Irrationality', Mark van Roojen effectively put the nail in the coffin of higher-order attitude approaches like Blackburn's, whatever their form, by showing that they *overgenerate* valid arguments, even if we grant that they work on their own terms. The easiest way to see the crux of van Roojen's point is to compare the following two arguments:

1a Stealing is wrong.
2a If stealing is wrong, then murder is wrong.
3a Murder is wrong.

1b Stealing is wrong.
2b It is wrong to both disapprove of stealing and not disapprove of murder.
3b Murder is wrong.

The problem is that on Blackburn's account, sentences 2a and 2b express the very same attitude – disapproval of the state of both disapproving of

stealing and not disapproving of murder. So if his account *does* suffice to explain why the former argument is valid, then it also suffices to explain why the latter argument is valid. But the latter argument is not, intuitively, valid. So Blackburn's higher-order attitudes approach overgenerates validity.

The problem, as van Roojen put it, is that the *kind* of rational incoherence that is generated by Blackburn's explanation is not the incoherence of having inconsistent beliefs or of having a belief and failing to draw one of its consequences. It is the kind of incoherence involved in thinking that murder is wrong and murdering anyway. He called this broader kind of incoherence – the kind that is too broad to generate an adequate account of validity – *Moorean* incoherence, after the oddity that Moore diagnosed in both believing that *p* and believing that one does not believe that *p*. The lesson of van Roojen's article is the main lesson we have learned from higher-order attitude accounts more generally: if expressivists are to be able to explain validity, they are going to need to appeal to a kind of incoherence among attitudes that is of a more specific type than the broad kind of incoherence to which Blackburn initially appealed. They are going to have to appeal to incoherence among attitudes that is of the very same type as the incoherence involved in both believing that *p* and also believing that ~*p*. This is what all expressivist approaches to the Frege-Geach Problem since the late 1980s have really been trying to do, with greater levels both of sophistication and of appreciation of the nature of the problem, over the ensuing years.

The Negation Problem

Since the mid-1990s, conditionals have attracted much less direct attention, and more attention has been paid to the case of negation. The reason for this is simple; at a minimum, explaining why conditionals validate *modus ponens* requires explaining why {'P','P→Q','~Q'} is an inconsistent set of sentences. But this problem has many moving pieces: it requires having in hand an expressivist account not only of the semantics of the conditional, but of negation, and an expressivist account of the inconsistency of sentences, besides.[3] So much investigation has adopted a more conservative strategy, and focused on trying to acquire an adequate expressivist semantics for negation first, so that it can be used as a *fixed point*, in developing an expressivist semantics for conditionals. The most important semantic property of negation, after all, is that negated sentences should turn out to be inconsistent with the sentences they negate, and to explain why {'P','~P'} is an inconsistent set, we need only know the semantics for negation and how inconsistency works – we don't need to know anything about conditionals.

Moreover, since van Roojen's article it has been fairly well understood that not just any old kind of mental incoherence or rational tension between two mental states will suffice in order to explain inconsistency

between the sentences that express them. The way that beliefs with inconsistent contents clash with one another is fine, but the way that having an attitude and disapproving of oneself for having that attitude clash is not fine. Allan Gibbard has coined a special technical term for mental states which clash in the sort of way that beliefs with inconsistent contents do – the right sort of way in order to explain inconsistency in the sentences which express them. He calls this kind of clash *disagreement*.

If ordinary descriptive beliefs were the only kinds of mental state that could disagree with one another, then it would follow immediately that the only way 'murder is wrong' and 'murder is not wrong' could be inconsistent, would be if they both express ordinary descriptive beliefs. But this is precisely what expressivists like Blackburn, Gibbard, and Horgan and Timmons deny.[4] But fortunately, as Gibbard has observed, following in the footsteps of Stevenson, beliefs are not the only kinds of mental states which appear to conflict with one another. As has received a great deal of attention in the philosophy of action, there is a very similar kind of rational conflict between *intending* inconsistent things as between *believing* inconsistent things.

So if intentions – or other noncognitive attitudes like *disapproval* – share with beliefs the property that they disagree with each other just in case they are toward inconsistent contents, then expressivists can hope to explain inconsistency between moral sentences and their negations, by assigning 'stealing is wrong' and 'stealing is not wrong' to states of disapproval of inconsistent things. For example, if 'stealing is wrong' expresses disapproval of stealing, and 'stealing is not wrong' expresses disapproval of not stealing, then we could use the fact that these two states disagree with one another in order to explain why 'stealing is wrong' and 'stealing is not wrong' are inconsistent.

But this approach meets an important obstacle. The obstacle is that even if disapproval of stealing and disapproval of not stealing disagree, the latter is not, in fact, the attitude expressed by 'stealing is not wrong', but rather that expressed by 'not stealing is wrong'. This means not only that we have failed to give an account of what mental state is expressed by 'stealing is not wrong', but also that there is no state that we can assign to it, such that we can explain all of the inconsistencies that we need to explain as cases of disapproving of inconsistent contents – which disagree in the same, non-Moorean way, as beliefs with inconsistent contents do: by being cases of the same attitude with inconsistent contents.

This is easy to prove. Compare the following four sentences:

1c Stealing is wrong. → DIS(stealing)
2c Stealing is not wrong. → DIS(x)
3c Not stealing is wrong. → DIS(not stealing)
4c Not stealing is not wrong. → DIS(y)

Both 1c and 2c are inconsistent sentences, as are 3c and 4c. So if their inconsistency is to be explained in terms of the non-Moorean disagreement between the mental states that they express – states which rationally conflict with each other in just the same way that beliefs with inconsistent contents do – and this is to be explained by the fact that disapproval, like belief and intention, is the sort of attitude that it rationally conflicts in this way to hold toward inconsistent contents, then 2b and 4c must express some states of disapproval. 2c must express disapproval of something inconsistent with stealing, in order to explain why 1c and 2c are inconsistent, and 4c must express disapproval of something inconsistent with not stealing, in order to explain why 3c and 4c are inconsistent. But if x is inconsistent with stealing, and y is inconsistent with not stealing, then it follows that x and y must be inconsistent with each other. But this yields the prediction that the states of mind expressed by 2c and 4c rationally conflict in exactly the way required in order to explain the inconsistency of 2c and 4c. But 2c and 4c are not inconsistent sentences!

The Hierarchy of Attitudes

Faced with this problem, which Simon Blackburn ('Attitudes and Contents') began to recognize and which has become increasingly vivid ever since, particularly thanks to a pair of important articles by Nicholas Unwin at the turn of the century, contemporary expressivist views have granted that 'stealing is not wrong' cannot express the same kind of attitude as 'stealing is wrong' expresses toward stealing. And this has led most theorists – including Blackburn, Gibbard, and Horgan and Timmons – to postulate that 'stealing is not wrong' expresses a new and *different* attitude toward stealing, which is nevertheless assumed to disagree with disapproval of stealing.[5] So rather than being inconsistent because they express the same attitude toward inconsistent contents, as 'grass is green' and 'grass is not green' are, these views hold that 'stealing is wrong' and 'stealing is not wrong' are inconsistent because they express different attitudes toward the *same* content – attitudes which just happen to disagree with one another.

This is the first step on the expressivist trip toward the postulation of an entire hierarchy of distinct noncognitive attitudes that can be expressed by moral sentences. Unwin used a simple example in order to illustrate where the pressure to postulate a new and distinct attitude expressed by the negations of atomic sentences comes from:

w Jon thinks that stealing is wrong.
n1 Jon doesn't think that stealing is wrong.
n2 Jon thinks that stealing is not wrong.
n3 Jon thinks that not stealing is wrong.

The task of providing an expressivist semantics for 'not' is the task of giving content to n2 – for expressivism is the view that you give the meaning of 'P' by saying what it is to think that P. But the trouble is that we can't just read this off of the expressivist account of the meaning of 'stealing is wrong', because w lacks sufficient structure. As n1 to n3 illustrate, there are three places in which a 'not' can be inserted in w. But as n1★ to n3★ below illustrate, there are *not* three places in which a 'not' can be inserted in the schematic expressivist account of w:

w★ Jon disapproves of stealing.
n1★ Jon doesn't disapprove of stealing.
n2★ ???
n3★ Jon disapproves of not stealing.

As Unwin's examples illustrate, the reason why expressivists have needed to resort to an attitude *distinct* from disapproval to be expressed by the negations of atomic 'wrong' sentences, is in order to make up for the lack of structure in their account of the attitude expressed by the atomic sentences. But this also suffices to show that the problem exists not only for negation, but for *every* complex-sentence-forming construction, as the following examples illustrate:

&1 Jon thinks that stealing is wrong and thinks that murdering is wrong.
&2 Jon thinks that stealing is wrong and murdering is wrong.
&3 Jon thinks that stealing and murdering is wrong.

∀1 Everything is such that Jon thinks that it is wrong.
∀2 Jon thinks that everything is wrong.
∀3 Jon thinks that (doing everything) is wrong.

P1 Jon thought that stealing is wrong.
P2 Jon thinks that stealing was wrong.
P3 Jon thinks that having stolen is wrong.

◇1 It is possible that Jon thinks stealing is wrong.
◇2 Jon thinks that it is possible that stealing is wrong.
◇3 Jon thinks that (possibly stealing) is wrong.

And it is easy to extend such examples indefinitely. For each case, all three sentences need to be distinguished. For each case, providing an expressivist semantics for that complex-sentence-forming construction is a matter of giving content to the second sentence. And for each case, there is one too few places in the structure of sentence w, for any account of the second sentence to fall out. So for each such construction, expressivists need to postulate a new attitude to be expressed by sentences formed by that construction, in order to make up for this lack of structure.

Moreover, things don't end there; for the same reasons that conjunctions of atomic sentences require a new attitude, conjunctions of negations of

atomic sentences require a new attitude, as do conjunctions of conjunctions. Similarly, for the same reasons that negations of atomic sentences require a new attitude, so do negations of conjunctions. So explicit advocates of the hierarchy of attitudes, like Horgan and Timmons ('Cognitivist Expressivism'), rapidly commit to thousands and thousands of distinct kinds of attitudes to be expressed by even relatively simple moral sentences.

Proponents of these kinds of expressivist views, like Gibbard (*Thinking How to Live*) and Horgan and Timmons ('Cognitivist Expressivism'), argue that everyone is in the same boat, and needs at some point to appeal to similar sorts of assumption. But their critics, including Schroeder (*Being For*), point out that descriptive sentences all express the same kind of attitude, rather than indefinitely many distinct kinds of attitude, and complain that appealing to a hierarchy of attitudes is like saying that complex sentences express that state of mind, whatever it is, that would ensure that they have the right semantic properties, rather than saying what that attitude is, and explaining why it has those properties, as ordinary descriptivist semantic theories can do.

Combining with Descriptive Language

In addition to the obstacles so far encountered, expressivists face a special problem in trying to provide a unified semantics for both moral and descriptive language. The problem, at bottom, is that in order to provide a semantics for 'not', 'if . . . then', and other constructions as they apply to *moral* sentences, expressivists acquire commitments which act as a *constraint* on their semantics for complex ordinary descriptive sentences like 'grass is not green'.

The basic problem is easy to state, even though exploring it in detail requires holding fixed a number of matters of detail about just how a given expressivist view works. The problem is that two-place connectives like 'and', 'or', and binary quantifiers can take two moral arguments, but they can also take one moral and one descriptive argument, or two descriptive arguments, as these examples illustrate:

1d Stealing is wrong or murder is wrong.
2d Stealing is wrong or grass is green.
3d Snow is white or grass is green.

If 'or' is to have the same meaning in all three of these sentences, therefore, then its meaning in 3d will be constrained by the commitments that expressivists need to adopt in order to get it to work in 1d and 2d. And the same goes, in principle, for one-place connectives, as well – if they are to have the same meaning in both moral and descriptive sentences, then their meaning in purely descriptive sentences will be constrained by the expressivist's commitments about how they need to work in moral sentences.

Bob Hale argued, on the basis of the approaches to disjunction advocated by Blackburn in the 1980s, that this problem saddled expressivists with uncomfortable results about ordinary descriptive disjunctions. Max Kölbel used this important fact to argue that expressivism is not, as some might have believed, a way to provide a nondescriptivist semantics for moral language while preserving an ordinary descriptivist semantics for non-moral language. Schroeder ('Expression for Expressivists') uses it in order to argue for important constraints on what 'express' could mean, within an expressivist theory, and in *Being For*, he argues that even the most promising sort of expressivist semantics will run afoul of this problem once it comes to accounting for more complicated and interesting linguistic constructions like tense, modals, and binary quantifiers. Cian Dorr, meanwhile, has used the case of mixed moral-descriptive sentences in order to raise new, epistemological problems for expressivism.

In general it is fair to say that dealing with the extra constraint that an expressivist treatment of moral language poses on how expressivists are to account for the meaning of ordinary complex descriptive sentences is one of the most central and difficult aspects of the Frege-Geach Problem. It is hard enough to give an adequate semantic account for a wide range of difficult-to-understand natural language constructions, from indicative conditionals to generics to epistemic modals. It's only harder to approach these problems under the kinds of extra constraints imposed by expressivism.

Is the Problem Really One about Truth or Logic?

Notice that the essence of the Frege-Geach Problem has nothing especially to do with truth, nor with logic, nor with reasoning. And it is certainly not simply a problem about explaining moral *modus ponens* arguments. Many authors *have* claimed that the Frege-Geach Problem is about accounting for logic, but we've seen here that that is far from the case. The problem has often been thought to be specifically about logic primarily because of the effectiveness of Geach's example of a *modus ponens* argument (which he used in order to argue that 'good' must mean the same thing when it appears inside the antecedent of a conditional as outside it) and because most discussions of the problem over the last twenty years have been heavily influenced by Blackburn's formulations, and he strongly emphasized this aspect of the problem.

But in fact the problem is much more general. Consider, for example, the case of attitude-ascriptions. 'Max hopes that this is good', 'Max wonders whether this is good', and 'Max is ecstatic that this is good' are just a small sampling of the very wide range of attitude-ascriptions. Each of these constructions works just like their counterparts with ordinary descriptive complements: 'Max hopes that this is green', 'Max wonders whether this is green', and 'Max is ecstatic that this is green'. Whereas the primary semantic properties of words like 'not', 'and', and 'if . . .

then' might be characterized as *logical* properties, the primary semantic properties of 'hopes that', 'wonders whether', and 'is ecstatic that' wouldn't be so characterized. Yet noncognitivists owe an account of the meaning of each and every attitude verb, just as much as they owe an account of 'not', 'and', and 'if . . . then'. Very little progress has yet been made on how noncognitivists can treat attitude verbs, and the prospects for further progress look dim.

In the 1990s, a number of authors suggested that the Frege-Geach Problem could be trivially solved by appeal to the idea that moral sentences can be true in a *minimal* sense. If 'it is true that this is good' means nothing over and above what 'this is good' means, these authors argued, then noncognitivists do not, after all, have any problem in accounting for how moral arguments can be valid. In an important paper from 1996, Jamie Dreier dismantled this idea, and explained why minimalism about truth did not, in fact, suffice to confront the issues that were actually at stake in the Frege-Geach Problem; I won't summarize Dreier's main points here, but rather make a more direct observation. Just note that nowhere in my statement of the Frege-Geach Problem and its history have I needed to use the words 'true' or 'false'. The Frege-Geach Problem is a problem quite independently of whether noncognitivists think that moral sentences can be true or false. That is as easy an answer as we might like for why minimalism about truth does not trivially suffice to solve the problem.

The idea that the problem could be solved by appeal to the idea that moral sentences can be true in a 'minimal' sense probably derives from a restricted conception of what the problem involves, in the first place. As I have characterized the problem, it is the problem of accounting for the meaning not only of atomic moral sentences, but of complex sentences containing moral terms – and explaining why moral *modus ponens* arguments are valid is only one small part of this task: it is part of establishing that one's semantics for conditionals suffices to predict and explain one of conditionals' most important semantic properties: that they license *modus ponens*. But if we instead characterized the problem in a restricted way, as the problem of explaining how moral arguments can be *valid*, and thought that this problem is hard because only sentences that can be true or false can figure in valid arguments, and noticed that noncognitivism has often been characterized as the view that moral sentences are *not* true or false, then we might, in fact, be led to think that minimalism about truth would suffice to solve this problem. So this is one important example of how a restricted conception of what the problem is has affected ideas about what it would take to solve it.

Hybrid Theories

In recent years, a number of philosophers have proposed that views which incorporate some expressivist elements and some cognitivist

elements could avoid or finesse the Frege-Geach Problem while still retaining other advantages of pure expressivism. These views are often inspired by the example of pejoratives, which appear to have both descriptive, truth-conditional, content, and also to involve some kind of attitudinal component (often a negative one). Despite their connection to negative attitudes, however, pejoratives fit smoothly into any and all complex linguistic constructions in exactly the same way that ordinary descriptive language does, so this licenses optimism that moral terms like 'wrong' might do so as well, if they differ from descriptive terms like 'green' in the same way that pejoratives like 'nigger' do.

It is still early to say what the final verdict will be on these hybrid approaches, but one early lesson is that hybrid theories differ widely over just how seriously they take the pejorative model. Daniel Boisvert's *Expressive-Assertivism* takes the model of pejoratives very seriously, and holds that moral terms, like pejoratives, have a fixed descriptive content and are associated with a fixed attitude – the same for every speaker. Competing hybrid theories like those of Michael Ridge and Stephen Barker, in contrast, hold that moral terms are quite different from pejoratives, varying from speaker to speaker in both their descriptive content and in which attitude they convey. Some authors, including van Roojen ('Expressivism, Supervenience'), have begun to explore the ways in which these differences open views like Ridge's and Barker's up to problems that views like Boisvert's do not face.

In the end, hybrid theories do not escape the burden of needing to supply a compositional semantics which explains why each sort of sentence has the right kinds of semantic properties. They merely provide some grounds for optimism – in proportion to how closely they are modeled on cases like pejoratives – that this is something any adequate semantics for English needs to be able to do, anyway. In doing so, they expand on a long-time noncognitivist strategy of looking to linguistic phenomena which any adequate natural language semantics will need to explain, anyway, both for optimism and for clues about how to develop a noncognitivist view.

Summary

At bottom, the essence of the Frege-Geach Problem is that moral and descriptive terms play exactly the same kind of semantic role in every kind of complex linguistic construction in natural languages. Since noncognitivist views consist centrally in the idea that moral terms like 'wrong' have a *different* kind of meaning from ordinary descriptive terms like 'green', that makes noncognitivism look like a very unpromising hypothesis about natural language semantics. At the very least, if noncognitivism is going to get off of the ground, noncognitivists have their work cut out for them – not only to explain the meaning of complex moral sentences, but to offer a compositional semantics for English that

predicts and explains why despite their differences, moral and descriptive terms function in all of the same kinds of ways. That is the Frege-Geach Problem for noncognitivist metaethical theories.

Acknowledgments

Special thanks to David Sobel and Mark van Roojen.

Short Biography

Mark Schroeder's research centers on issues in and around practical reason and metaethics, particularly including normative ethics, the philosophy of language, metaphysics, epistemology, and the history of ethics. He has published articles on these topics in *Ethics*, *Noûs*, *Philosophy and Phenomenological Research*, *Philosophical Studies*, *Philosophers' Imprint*, *Australasian Journal of Philosophy*, *Philosophical Perspectives*, *Journal of Ethics and Social Philosophy*, *Pacific Philosophical Quarterly*, *Oxford Studies in Metaethics*, *Social Theory and Practice*, and the *Stanford Encyclopedia of Philosophy*, as well as two books, *Slaves of the Passions* and *Being For: Evaluating the Semantic Program of Expressivism*, with Oxford University Press. *Being For* takes the first systematic look at what expressivists need to do in order to make progress in solving the Frege-Geach problem, and explores the commitments that they must take on to do so. He holds a Ph.D. in philosophy from Princeton University, has taught at the University of Maryland, and is currently associate professor of philosophy at the University of Southern California.

Notes

⋆ Correspondence address: School of Philosophy, University of Southern California, 3709 Trousdale Pkwy, Los Angeles, CA 90089, USA. Email: maschroe@usc.edu.

[1] Horgan and Timmons eschew the label, 'noncognitivism', for reasons that need not concern us here.
[2] Here I use 'accept' as a shorthand for 'is in the mental state expressed by'.
[3] Expressivists can't appeal directly to truth or satisfaction in their account of inconsistency of sentences, because their semantics doesn't generate truth-conditions; only mental states.
[4] Horgan and Timmons do not deny this, if the 'ordinary descriptive' qualification is removed.
[5] Strictly speaking, these observations only apply to the *basic* normative predicate appealed to by each theorist. For Gibbard (*Thinking How to Live*), this is 'the thing to do', for Horgan and Timmons ('Cognitivist Expressivism') it is 'ought', which they construe as a sentential operator; in general it is different for each theorist.

Works Cited

Alm, David. 'Moral Conditionals, Non-Cognitivism, and Meaning'. *The Southern Journal of Philosophy* 38.3 (2000): 355–77.
Ayer, A. J. *Language, Truth, and Logic*. New York, NY: Dover, 1936.

Björnsson, Gunnar. 'Why Emotivists Love Inconsistency'. *Philosophical Studies* 104.1 (2001): 81–108.
Blackburn, Simon. 'Attitudes and Contents'. *Ethics* 98.3 (1988): 501–17.
——. *Essays in Quasi-Realism*. Oxford: Oxford UP, 1993.
——. 'Moral Realism' (1973). Reprinted in Simon Blackburn. *Essays in Quasi-Realism*. Oxford: Oxford UP, 1993. 111–29.
——. *Ruling Passions*. Oxford: Oxford UP, 1998.
——. *Spreading the Word*. Oxford: Oxford UP, 1984.
Boisvert, Daniel. 'Expressive-Assertivism'. *Pacific Philosophical Quarterly*, forthcoming.
Carnap, Rudolf. *Philosophy and Logical Syntax*. Bristol, UK: Thoemmes Press, 1996 [1935].
Copp, David. 'Realist-Expressivism: A Neglected Option for Moral Realism'. *Social Philosophy and Policy* 18 (2001): 1–43.
Cuneo, Terrence. 'Saying What We Mean: An Argument Against Expressivism'. *Oxford Studies in Metaethics* 1 (2006): 35–71.
Dorr, Cian. 'Non-Cognitivism and Wishful Thinking'. *Noûs* 36.1 (2002): 97–103.
Dreier, James. 'Expressivist Embeddings and Minimalist Truth'. *Philosophical Studies* 83.1 (1996): 29–51.
——. 'Negation for Expressivists: A Collection of Problems with a Suggestion for their Solution'. *Oxford Studies in Metaethics, Vol. 1*. Oxford: Oxford UP, 2006. 217–33.
Enoch, David. 'How Noncognitivists Can Avoid Wishful Thinking'. *Southern Journal of Philosophy* 41 (2003): 527–45.
Geach, P. T. 'Ascriptivism'. *Philosophical Review* 69 (1960): 221–5.
——. 'Assertion'. *Philosophical Review* 74 (1965): 449–65.
——. 'Imperative and Deontic Logic'. *Analysis* 18 (1958): 49–56.
Gibbard, Allan. *Thinking How to Live*. Cambridge, MA: Harvard UP, 2003.
——. *Wise Choices, Apt Feelings*. Cambridge, MA: Harvard UP, 1990.
Hale, Bob. 'Can There Be a Logic of Attitudes?' *Reality, Representation, and Projection*. Eds. J. Haldane and C. Wright. New York, NY: Oxford UP, 1993. 337–64.
Hare, R. M. *The Language of Morals*. Oxford: Oxford UP, 1952.
——. 'Meaning and Speech Acts'. *The Philosophical Review* 79.1 (1970): 3–24.
Horgan, Terry and Mark Timmons. 'Cognitivist Expressivism'. *Metaethics after Moore*. Eds. T. Horgan and M. Timmons. Oxford: Oxford UP, 2006. 255–98.
——. 'Nondescriptivist Cognitivism: Framework for a New Metaethic'. *Philosophical Papers* 29 (2000): 121–53.
Horwich, Paul. 'Gibbard's Theory of Norms'. *Philosophy and Public Affairs* 22 (1993): 67–78.
Kölbel, Max. *Truth without Objectivity*. New York, NY: Routledge, 2002.
Price, Huw. 'Semantic Deflationism and the Frege Point'. *Foundations of Speech Act Theory: Philosophical and Linguistic Perspectives*. Ed. S. L. Tsohatzidis. London: Routledge, 1994. 132–55.
Ridge, Michael. 'Ecumenical Expressivism: Finessing Frege'. *Ethics* 116.2 (2006): 302–36.
van Roojen, Mark. 'Expressivism and Irrationality'. *The Philosophical Review* 105.3 (1996): 311–35.
——. 'Expressivism, Supervenience, and Logic'. *Ratio* n.s. 18.2 (2005): 190–205.
Schroeder, Mark. *Being For: Evaluating the Semantic Program of Expressivism*. Oxford: Oxford UP, 2008.
——. 'Expression for Expressivists'. *Philosophy and Phenomenological Research*, forthcoming 2008.
——. 'How Expressivists Can and Should Solve Their Problem with Negation'. *Noûs*, forthcoming 2008.
Schueler, G. F. 'Modus Ponens and Moral Realism'. *Ethics* 98.3 (1988): 492–500.
Searle, John. 'Meaning and Speech Acts'. *Philosophical Review* 71 (1962): 423–32.
Sinnott-Armstrong, Walter. 'Expressivism and Embedding'. *Philosophy and Phenomenological Research* 61.3 (2000): 677–93.
Smart, J. J. C. *Ethics, Persuasion, and Truth*. London: Routledge and Kegan Paul, 1984.
Stevenson, C. L. 'The Emotive Meaning of Ethical Terms' (1937). Reprinted in C. L. Stevenson. *Facts and Values*. Westport, CT: Greenwood Press, 1963. 10–31.
——. *Ethics and Language*. Oxford: Oxford UP, 1944.
——. *Facts and Values*. Westport, CT: Greenwood Press, 1963.

Stoljar, Daniel. 'Emotivism and Truth Conditions'. *Philosophical Studies* 70 (1993): 81–101.
Unwin, Nicholas. 'Norms and Negation: A Problem for Gibbard's Logic'. *The Philosophical Quarterly* 51.202 (2001): 60–75.
——. '*Quasi*-Realism, Negation and the Frege-Geach Problem'. *The Philosophical Quarterly* 49.196 (1999): 337–52.
Zangwill, Nick. 'Moral Modus Ponens'. *Ratio* n.s. 5.2 (1992): 177–93.