philosophy of language for metaethics
forthcoming in the Routledge Companion to the Philosophy of Language

Metaethics is the study of metaphysics, epistemology, the philosophy of mind, and the philosophy of language, insofar as they relate to the subject matter of moral or, more broadly, normative discourse – the subject matter of what is good, bad, right or wrong, just, reasonable, rational, what we must or ought to do, or otherwise. But out of these four ‘core’ areas of philosophy, it is plausibly the philosophy of language that is most central to metaethics – and not simply because ‘metaethics’ was for a long time construed more narrowly as a name for the study of moral language. The philosophy of language is central to metaethics because both the advantages of and the open problems facing different metaethical theories differ sharply over the answers those theories give to central questions in the philosophy of language. In fact, among the open problems over which such theories differ, are included particularly further problems in the philosophy of language. This article briefly surveys a range of broad categories of views in metaethics and both catalogues some of the principal issues faced by each in the philosophy of language, as well as how those arise out of their answers to more basic questions in the philosophy of language. I make no claim to completeness, only to raising a variety of important issues.

1 cognitivist-realist theories

Probably the default view in metaethics is to assume that moral language does not differ in any deep or important way from non-normative language, so that sentences like ‘stealing is wrong’ or ‘it’s good that M ax is happy’, like sentences like ‘stealing is common’ or ‘it seems that M ax is happy’ are about some subject matter in the world – what is wrong or good. On the simplest version of such a view, ‘it’s good that M ax is happy’ expresses a single, invariant proposition across all contexts of utterance, but much the same issues face more careful views according to which all context-dependence of ‘it’s good that M ax is happy’, if there is any, is of an ‘unsurprising’ sort – due to facts such as that ‘good’, like ‘tall’, is a gradable adjective, or due to its tense.
One of the primary issues facing such metaethical theories in the philosophy of language concerns how we manage to talk about normative properties or facts. Consequently, metaethical realists are particularly interested in, or at least ought to be particularly interested in, the theory of reference - metasemantics, as it is sometimes called. The issues of importance to metaethical realists in the theory of reference, however, depend on the realist's answer to the metaphysical question of just what normative sentences are about. According to the nonreductive metaethical realist, normative sentences are about a sui generis normative domain, separate from any natural fact investigatable by the rest of the natural or social sciences. Consequently, nonreductive metaethical realists have a large stake in the debate between causal and inferential-role theories of reference - if sui generis normative facts are causally inert, then a causal theory of reference will not enable us to think or talk about them. This drives some work by metaethical realists outside of the philosophy of language - to account for how nonreductive, sui generis normative facts could in fact play a causal role (see particularly Sturgeon [1985], [2006] and Thom son [1996]) - and it motivates a general interest in non-causal theories of reference, for those who don’t accept this conclusion (see particularly Wed gwood [2007]).

Reductive metaethical realists, in contrast, hold that normative properties or facts can be analyzed in terms of non-normative or natural properties or facts - so that the subject matter of normative sentences is part of the world revealed to us by the natural and social sciences. Reductive realists' stake in the theory of reference is different: they must be able to explain why the analyses that they provide are surprising. Reductive metaethical realists therefore have a deep stake in understanding philosophical analysis in general and in the philosophy of language underlying surprising property identities, in particular. Much of metaethical inquiry has been concerned with this set of issues, and the most extensive work by metaethical realists in the philosophy of language has been devoted to it (see particularly Moore [1903], Boyd [1988], Brink [1989], Railton [1986], Smith [1994], and Jackson [1998]).

Reductive realists have been quick to point out that philosophy is full of surprising identities - including, the identity of water with H₂O. But things are not quite so easy. It is familiar how the identity of water and H₂O could be surprising - because it is an empirical finding - and familiar how it could be an empirical finding - because ‘water’ could have referred to something else, if our environment had been different, even though people took it to have the same inferential significance. But it is far from clear that ‘wrong’ could easily have referred to something else, simply because our environment had been different and even though people took it to have the same inferential significance - on the contrary, if we imagine cases in which people call different things wrong merely because of a difference in their environment, it is
fairly easy to pump the intuition that we simply disagree with these people about what is wrong. (See particularly Hare [1952], Dreier [1990], and Horgan and Timmons [1991], [1992].)

The apparent non-'twin-earthability' of 'wrong' raises general issues in the philosophy of language about whether surprising analyses like the claim that water is H₂O must be a posteriori (or might be synthetic without being a posteriori) and about whether making an a posteriori discovery always involves ruling out some genuine metaphysical possibility – as discovering that water is H₂O involves ruling out the possibility that we are on twin earth. This has interested metaethicists in whether other terms – for example, for functional kinds – exhibit similarly twin-earth resistant behavior, and in examples of difficult-to-discover but non-a posteriori analyses outside of metaethics – for example, the mathematical definition of continuity (see particularly Copp [2000], van Roojen [2006], Schroeder [2007b]). It also raises general questions from the philosophy of language about how people could succeed at talking about the same thing as one another, even while having very different ideas about the extension of those terms (see particularly Horgan and Timmons [2000], Dreier [1993]).

2 contextualist theories

As suggested earlier, context-dependence in moral sentences that is due to unsurprising features such as tense or the presence of gradable adjectives is something for which ordinary realist theories in metaethics can allow. But some theories – interestingly contextualist theories – hold not only that 'it is good that Max is happy' expresses different propositions in different contexts of utterance, but that it does so even when controlling for such unsurprising sorts of context sensitivity. It is possible, of course, for a non-reductive metaethical realist to also endorse this kind of surprising contextualism, but in practice, surprising contextualist theories are used primarily to make it easier to locate normative facts in the natural world. On such views, an important part of why it does not seem to be possible to identify moral facts with any particular natural facts is that moral claims express different natural facts, in different contexts of utterance (see particularly Dreier [1990], Timmons [1999], and Finlay [2004], [2009]).

The traditional form taken by metaethical contextualism is speaker subjectivism. The most naïve forms of speaker subjectivism identify the meanings of 'stealing is wrong' and 'I disapprove of stealing' for some attitude verb like 'disapprove', which picks out a negative desire-like attitude. More cautious forms identify not the meaning of these two expressions, but their truth relative to contexts of utterance and assignments (see Ayer [1936] for an argument which motivates this move). More cautious forms yet allow that there may not be any word in ordinary English for the requisite desire-like attitude, so 'disapprove'
needs to be understood as a stipulative technical term (see Urmson [1968] for an argument that motivates this move). Contemporary twists on speaker subjectivism include actually-rigidified speaker subjectivism, and speaker-group or conversational subjectivism, which appeals to something more like ‘we disapprove of stealing’ (for both, see Dreier [1990]).

Another contemporary twist on speaker subjectivism is to hold that there is a function from the speaker’s attitudes to the content of ‘stealing is wrong’ but that these contents are not themselves about the speaker’s attitudes. This view can be thought of as identifying the contents of ‘stealing is wrong’ and something like ‘stealing instantiates dthat[the property I disapprove of actions for instantiating]’ (compare Barker [2002], Schroeder [2009]). Even more flexible views require that the context contribute some parameter, and allow that in some contexts this parameter is fixed by the speaker’s attitudes or interests, but also allow for ‘disengaged’ contexts in which the parameter can be fixed by other people’s attitudes or interests, in the way that it is plausible that uses of ‘local’ refer by default to the neighborhood of the speaker’s location, but can also easily refer to the neighborhood of some other location (see Timmons [1999], Finlay [2009]).

Contextualists in metaethics face a set of problems familiar from other areas of philosophy. In particular, natural tests for context-dependence, including the behavior of cross-contextual indirect discourse reports and truth ascriptions, seem to yield the wrong results for many normative sentences. For example, consider the following bit of reasoning:

Immanuel (in one context): Lying is always wrong.
Jeremy (in a different context): Immanuel just said that lying is always wrong.
Jeremy: But lying is not always wrong.
Jeremy: So Immanuel just said something false.

On the face of it, Jeremy’s reasoning in this scenario makes perfect sense. But from the contextualist’s point of view, it is very puzzling. For either Jeremy is correctly reporting what Immanuel said or he is making an error. If he is correctly reporting what Immanuel said, then since Immanuel’s context is different and hence he did not in fact say anything inconsistent with what Jeremy means by going on to say that lying is always wrong, it would seem to follow that Jeremy’s inference is invalid. But it seems both that Jeremy can correctly report what Immanuel said in this way, and that he can legitimately go on to reason in this way – at one and the same time. Sophisticated contextualist views can account for the correctness of the indirect discourse report (in some contexts), and can account for the validity of the
inference (in other contexts), but the hard thing is accounting for both at the same time (and in the same context - compare Schroeder [2009]).

Another important and related issue facing contextualists is how to explain how speakers are able to engage in moral disagreements with one another. This problem was particularly acute for speaker-subjectivists, for whom even speakers in the same conversation were in relevantly different contexts; focusing instead on the speaker’s conversational partners allows people in the same conversation to disagree but still runs into challenges explaining disagreement across conversations (compare Dreier [1990]).

3 relativism
An alternative to metaethical contextualism is genuine relativism. According to genuine relativism, though ‘stealing is wrong’ expresses the same proposition in different contexts of utterance (again controlling for possible effects of less controversial context-dependent features), that proposition is not true or false simpliciter, but may be assessed differently as true or false from different perspectives (see MacFarlane [2005], [2007]). Like metaethical contextualism, relativism is compatible with nonreductive views in metaethics, but it is hard to see what the motivation for such a position would be; relativism is more naturally motivated on similar grounds to those for contextualism - the relativist may claim that the truth of a normative proposition relative to a given point of view is itself a straightforward natural matter of fact.

As a matter of the history of the literature, metaethicists have been slow to embrace and articulate genuinely relativist views; despite the historical importance of moral relativism, many careful discussions of relativism in metaethics interpret it as a kind of contextualism (compare Harman [1975], [1978], Wong [1984]). This is in part due to the fact that the resources for articulating relativism in clearly coherent ways are a more recent development in the philosophy of language. Given the historical importance of relativism in metaethics - dating at least to the presocratics - the philosophical issues surrounding relativism are particularly important for metaethics.

4 noncognitivist theories: expressivism
Probably the most interesting issues arising for metaethics in the philosophy of language arise for metaethical theories which depart quite radically from conventional ideas about semantic content. These include all of the theories in the ‘noncognitivist’ tradition exemplified by Hägerström, Ogden and Richards, Schlick and Carnap, Ayer and Stevenson, Hare, Blackburn, and Gibbard. The most developed contemporary strand of this tradition is the contemporary theory of metaethical expressivism, according to
which rather than understanding the meaning of normative sentences in terms of what what propositions they express or under what conditions they are true, we should understand the meaning of normative sentences in terms of what it is to have the corresponding normative thoughts (see especially Blackburn [1993] and Gibbard [1990] and [2003]). From the expressivist’s point of view, speaker subjectivists were on the right track to think that the right way to understand the meaning of ‘stealing is wrong’ is to compare it to ‘I disapprove of stealing’, but speaker subjectivists went too far to identify the meaning or even the truth conditions of these two sentences.

Compare, for example, what would happen if we identified the meanings of ‘grass is green’ and ‘I believe that grass is green’. It follows from that view that anyone who believes that grass is green and says ‘grass is green’ speaks truly. This is analogous to the prediction of speaker subjectivism that anyone who disapproves of stealing and says ‘stealing is wrong’ speaks truly. Moreover, this view predicts that ‘if I didn’t believe that grass is green, then grass wouldn’t be green’ expresses a truth in the mouth of every speaker; this is analogous to the prediction of speaker subjectivism that ‘if I didn’t disapprove of stealing, then it wouldn’t be wrong’ expresses a truth in the mouth of every speaker. So expressivists reason that speaker subjectivists were making the very same mistake as this view about the meaning of ‘grass is green’ does. Since we all know that the right relationship between ‘grass is green’ and ‘I believe that grass is green’ does not lead to these problems, expressivists propose to postulate that relationship between ‘stealing is wrong’ and ‘I disapprove of stealing’. We can stipulatively say that the first member of each pair expresses the attitude that the second member reports (see Schroeder [2008], Gibbard [2003]).

This view quickly lands expressivists in the midst of needing to answer very general questions in the philosophy of language about what the relationship is between ‘grass is green’ and the belief that grass is green. To answer these questions, expressivists need to take stands about the nature and norms on assertion and on the primacy of mental versus linguistic content. It also leads expressivists to a very different conception of the primary goals of semantic inquiry – whereas conventional theories take the goal of a semantic theory to be to assign truth values or at least truth conditions to sentences, expressivists are led to the idea that a semantic theory needs to work by associating each sentence, ‘P’, with what it is to think that P. His novel expressivist approach to the objectives and structure of semantic theorizing incurs deep commitments for expressivists, and the projects of justifying it and of carrying it out go to the heart of the semantic enterprise.

One potential source of solace for metaethical expressivists is the potential for applications of a similar semantic framework outside of metaethics. For example, conditionals, probability judgments, epistemic modals, and truth have all been motivated as calling for a distinctively expressivist treatment (see
Adams [1975], Gibbard [1981], Price [1983], Yalcin [forthcoming], Schroeder [forthcoming]. On such a view, conditionals express, rather than report, the conditional confidence of the speaker in the consequent, conditional on the antecedent; probability judgments and epistemic modals express, rather than report, the speaker’s credences; and truth and falsity ascriptions express, rather than report, the speaker’s agreement or disagreement with the object of the ascription. The better motivated such further applications for expressivism, the more optimism it makes sense to have that something like the expressivist’s approach to semantic theory can be carried out – so proponents and opponents of metaethical expressivism alike have a stake in all of these topics.

5 noncognitivist theories: illocutionary theories

Whereas expressivism seeks to account for the meaning of normative sentences by accounting for the nature of normative thought, an earlier generation of noncognitivist theories sought to account for the meaning of normative sentences by accounting for the nature of normative assertion. These broadly illocutionary theories seek to explain the meaning of a sentence, ‘P’, by saying what ‘P’ is used to do. Like expressivism, these theories depart from more conventional semantic approaches by declining to tell us what normative sentences are about, or what sort of state of the world would make them true – what is significant about the meaning of normative sentences, on such views, is that they are used to commend, to cajole, to discourage, or to incite – for example (see particularly Ayer [1936], Stevenson [1937], [1944], and Hare [1952]).

Illocutionary theories of normative language are obviously deeply involved in needing to resolve a wide range of issues from speech act theory; somewhat less obvious but equally important is the interest to speech act theories of an adequate understanding of the semantics of mood. Sentential mood is, after all, the conventionally-mediated aspect of linguistic use that is most transparently connected with illocutionary force – though declaratives can be used to ask questions and interrogatives can be used to issue commands and imperatives can be used to make assertions, declarative mood is conventionally associated with assertion, interrogative mood is conventionally associated with questioning, and imperative mood is conventionally associated with commands or requests. Illocutionarily-based metaethical theories take these facts as grounds for optimism that a semantic theory must ultimately be able to do the kinds of things that they ask of it (see Hare [1952], Smart [1984]).

Unfortunately, it isn’t immediately obvious that a general theory of mood must be able to do the things that metaethical illocutionary theories require of it – for example, on many conventional approaches
to mood, it is a feature only of whole sentences, and hence doesn’t embed. So, although there are complex imperative sentences like ‘don’t steal’ and ‘steal everything you can’, these are not cases in which imperatives embed in more complex constructions, but in which the imperative mood takes scope over a complex sentence–in other words, ‘don’t steal’ has the structure, !!(~(you steal)), rather than the structure, ~(!!(you steal)). But importantly, normative predicates do embed in complex constructions. So if mood doesn’t embed, mood is not a very useful comparison for illocutionary metaethical theorists.

However, there is some evidence that mood can embed in complex constructions. For example, there is some evidence that conditional interrogatives like ‘if it’s raining, then why am I not wet?’ are not interrogatives about conditionals, and that conditional imperatives like ‘if you’re going to the store, then bring me some milk’ do not issue commands to do conditional things (compare Barker [2004]). Moreover, though it is difficult to form conjunctions of sentences of different moods using ‘and’, sentences like ‘close the window, for it’s freezing in here’ and ‘it’s cold in here, but leave the window open’ make perfect sense and do seem to combine conjuncts of imperative and declarative moods. Advocates of illocutionary metaethical theories are wise to be particularly interested in such cases, and in whether they put pressure on an adequate semantic theory of mood to allow for ways in which conventional markers of illocutionary force like mood can in fact compose into logically more complex constructions–as normative predicates need to be able to do (compare especially Boisvert and Ludwig [2006]).

6 general issues facing noncognitivist theories

In addition to the issues in the philosophy of language of special interest to expressivists, and those of special interest to illocutionary theories, both strategies for implementing the general noncognitivist program in metaethics face a shared set of difficult issues in the philosophy of language, centered around the famous Frege-Geach Problem.

The core of the Frege-Geach Problem is that of how to account for compositionality. The central problem is that since these theories are non-extensional, they can’t rely on the feature of broadly extensional theories of meaning, that negation corresponds to complementation, conjunction to intersection, and disjunction to union. This is because lacking the state of mind expressed by ‘P’ is not the same as having the state of mind expressed by ‘~P’ and having both the states of mind expressed by ‘P’ and by ‘Q’ is not the same as having the state of mind expressed by ‘P&Q’. Similarly (for illocutionary theories), performing the illocutionary act associated with ‘~P’ is clearly different from simply not performing the illocutionary act associated with ‘P’, and performing the illocutionary act associated with ‘P or Q’ is clearly different
from either performing the illocutionary act associated with ‘P’ or performing that associated with ‘Q’. So if a noncognitivist theory of meaning is to explain what it is to think that ~P on the basis of what it is to think that P, or to explain what illocutionary act ‘~P’ is suited to perform on the basis of what illocutionary act ‘P’ is suited to perform, the task is far from straightforward (see especially Unwin [1999], [2001] and Schroeder [2008]; see Geach [1965] for the version of this problem which gave it its name).

Moreover, all noncognitivist metaethical theories need to confront very general issues in the philosophy of logic. Whatever meaning is associated with ‘not’, it had better suffice to explain why for any sentence ‘P’, ‘P’ and ‘it is not the case that P’ cannot both be true. Consequently, they must either take a very indirect approach of trying to earn their way to this very simple result, or they must construct an alternative entire theory of inconsistency – and relatedly, of validity and logical consequence (see particularly Blackburn [1984], Gibbard [1990], [2003], and Schroeder [2008]).

Both varieties of noncognitivist theory are also deeply committed to understanding the nature of truth. According to early proponents of noncognitivist theories, normative sentences cannot be true or false, but contemporary proponents have sought to explain how they can be true or false, by appeal to broadly deflationary accounts of truth (see especially Blackburn [1993]). This is an important move, not only because we do in ordinary discourse attribute truth or falsity to normative claims, but in order to make good on the logical relationships between normative sentences on anything like ordinary views about what logical relationships involve – even by a very indirect route. Noncognitivists are also committed to solving problems about truth-aptness that do not face other theories; for example, a noncognitivist who takes expressives like ‘dammit, not stealing!’ as a semantic model for the meaning of ‘stealing is wrong’ needs to explain why the latter is truth-apt even though the former is not (compare Dreier [1996]).

Noncognitivists also face special challenges in accounting for the semantics of attitude ascriptions and indirect discourse reports – sentences like ‘Max thinks that stealing is wrong’ and ‘Max said that stealing is wrong’. Indeed, given the special constraints under which noncognitivist semantic theories operate, there are pressing questions, even if they can account for the simple compositionality of the ‘truth-conditional’ connectives, of whether they will be able to account for more general compositional issues – including fitting in with generalized quantifiers, generics, modals, tense, and mood. Since noncognitivist theories really propose a quite different way of thinking about the goals of a semantic theory, their ultimate debt is to be able to carry such a semantic program out across the board, which means that they essentially
have commitments for understanding the semantics of every part of natural languages (see particularly Schroeder [2008]). Commitments for noncognitivism in the philosophy of language run deep.

7 hybrid theories

Some metaethical theories hold that it is not sufficient to understand the meaning of normative sentences to focus either on what such sentences are about or on what such sentences are used to do - for the meaning of normative sentences involves both a ‘descriptive’ content and an expressive or illocutionary component. Whereas expressivists focus on the nature of assertion in general or on expressives and illocutionarily-oriented noncognitivists focus on imperatives and performatives, hybrid theorists typically take as their examples racial slur terms, and sometimes other ‘thick’ evaluatives. Racial slurs seem to many to be transparently associated both with a particular descriptive content (that of belonging to a certain race) and to be conventionally associated with expressing contempt or denigrating (no coincidence of etymology, here). It is the position of hybrid views that normative words are at least this much like racial slurs. (See particularly Copp [2001], Barker [2000], Boisvert [2008], and Ridge [2006], [2007].)

In fact, however, theories about the semantics of racial slurs vary widely, and so hybrid theorists are well-advised to pay attention not only to slurs, but also to pejoratives and epithets like ‘jerk’, ‘asshole’, and ‘bastard’, each of which also seems to many to be associated with some descriptive content as well as a derogatory function, and to paradigmatic ‘thick’ evaluatives like ‘courageous’ and ‘kind’. Another good model for hybrid theorists to be interested in is the phenomenon of intonational sneer – of the kind associated with the claim not simply that someone is fat, but that he is ‘FA-at’ (compare Blackburn [1992]). Though intonation is a feature of language that it is easy to forget about when we spend most of our time as theorists comparing representations of sentences on the written page, it is very plausible that intonational sneer is conventionally determined – it works even with unnatural words, for example, in the claim that someone is ‘SMA-art’, in the same tone of voice. The diversity of features involved with racial slurs, pejoratives and other epithets, ‘thick’ evaluative terms, and intonational sneer gives hybrid theorists a variety of models to work with, in seeking to understand normative language, but each of these topics needs to be much better understood in its own right, before any hybrid theory is properly up and running.

One of the most important things hybrid theorists need to understand about normative language, as well as about each of these examples from other parts of language, is how the illocutionary or expressive component of these terms’ meaning projects through environments like negation, the antecedents of conditionals, attitude ascriptions, and indirect-discourse reports. A variety of the payoffs for hybrid
theories turn out to turn on the hypothesis that the expressive component of normative language projects through negation and the antecedents of conditionals, but not through belief ascriptions - but it’s far from clear that racial slurs work in this way, so much remains for hybrid theorist to sort out (see Schroeder [2009], Potts [2003]).

8 general issues for metaethics from the philosophy of language

In addition to all of the specific issues catalogued above facing metaethicists in virtue of the specific metaethical theories which they adopt, there are general issues related to language of broad interest to all views in metaethics. For example, any theory about ‘good’ or the nature of the property it expresses - whether realist or expressivist - needs to start from a proper understanding of the argument structure of ‘good’ - whether it expresses a monadic property of states of affairs, for example, or a relation between states of affairs and agents whom those states of affairs are good for, or a function from properties to properties, or something else altogether (compare Ziff [1960], Geach [1956], Thomson [2003], Kraut [2007]). Similarly, any theory about ‘ought’ - whether realist or expressivist - needs to start from a proper understanding of whether it is a context-independent propositional operator, a context-dependent propositional operator (and if so what arguments are provided by context), or whether it expresses a relation between agents and actions (compare Wedgwood [2007], Schroeder [unpublished], Cariani [2009]). And arguably substantive issues in both normative and metaethics can also turn on the argument structure of claims about reasons (see Schroeder [2007a]). Answering all of these questions requires addressing general issues from both linguistics and the philosophy of language.

One particular issue of general interest to metaethicists is the relationship between epistemic and deontic modals. Early treatments of deontic logic were inspired by analogies with epistemic and alethic modalities, and both cross-linguistic data and the fairly pervasive availability of both epistemic and deontic readings for the very same modals (‘must’, ‘have to’, ‘may’, ‘ought’, ‘should’) encourages the hypothesis that these terms make a uniform semantic contribution to both epistemic and deontic readings (see von W right [1951], K ratzer [1977], [1981]). Depending on how this uniform semantic contribution is made out, it potentially constitutes a considerable challenge to metaethicists who think that normative or deontic language is special (see especially Finlay [2009]).

As I’ve been indicating in this article, the issues connected to metaethics from the philosophy of language are extremely diverse; I’ve had almost nothing to say about any of them here, only an extreme
bird’s-eye glance at the sweep of the territory. But it’s safe to say that there is much work for philosophers of language to do in metaethics, and much work for metaethicists to do, in the philosophy of language.¹

references


¹ Special thanks to Gillian Russell and Delia Graff Fara.


[forthcoming]. 'Oughts and Ends.' Forthcoming in Philosophical Studies.

Geach, Peter [1956]. 'Good and Evil.' Analysis 17: 33-42.


______ [forthcoming]. 'How to be an Expressivist About Truth.' Forthcoming in New Waves in Truth, edited by Cory Wright and Nicholas Jang Peterson.


